

124425

STIC-Biotech/ChemLib

From: Mertz, Prema
Sent: Sunday, June 13, 2004 11:56 AM
To: STIC-Biotech/ChemLib
Subject: 10/009,557

Please search SEQ ID NO:34 with DNA databases.

Please search SEQ ID NO:9 with protein databases.

Thanks.

Prema Mertz, Ph.D.
Primary Examiner
Art Unit 1646
4D81 Remsen Bldg Mailbox 4C70
US Patent & Trademark Office
(571) 272-0876

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 6/16
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2004, 09:58:40 ; Search time 45 Seconds
(without alignments)
2369.892 Million cell updates/sec

Title: US-10-009-557-9

Perfect score: 1745

Sequence: 1 MYVLSPEVFIIQLLFIQAI.....PIISYFCSLGCYVNSSDMLK 338

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTRMBL.25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodeo.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvivirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	284.5	16.3	428	5 Q9VKU0	Q9VKU0 drosophila
2	283.5	16.2	428	5 Q24386	Q24386 drosophila
3	247	14.2	366	11 Q8BX06	Q8BX06 mus musculus
4	247	14.2	403	4 Q96CX6	Q96CX6 homo sapien
5	244	14.0	311	16 Q8F116	Q8F116 leptospira
6	233.5	13.4	498	16 Q8F3F9	Q8F3F9 leptospira
7	227	13.0	423	16 Q8F7S1	Q8F7S1 leptospira
8	225.5	12.9	507	5 Q9N3F2	Q9N3F2 caenorhabdi
9	225	12.9	633	16 Q8F3F6	Q8F3F6 leptospira
10	224.5	12.9	452	16 Q8F118	Q8F118 leptospira
11	222.5	12.8	428	16 Q8F3F8	Q8F3F8 leptospira
12	218	12.5	679	5 Q61967	Q61967 caenorhabdi
13	216	12.4	426	16 Q8F213	Q8F213 leptospira
14	215	12.3	289	13 Q7ZSY7	Q7ZSY7 xenopus lae
15	212.5	12.2	1537	4 Q96NW7	Q96NW7 homo sapien
16	211.5	12.1	377	16 Q8F119	Q8F119 leptospira

ALIGNMENTS

RESULT 1

ID	Q9VKU0	PRELIMINARY;	PRT;	428 AA.
AC	Q9VKU0;			
DT	01-MAY-2000 (TrEMBLrel. 13, Created)			
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)			
DE	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)			
DE	IRR47 protein (GH06740P).			
GN	IRR47 OR C6098.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Berkely;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.P.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mikos S.L.G.,			
RA	Abriel J.P., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,			
RA	Ballev R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.X.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Buck J., Brokstein P., Brottier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,			
RA	de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.N.,			
RA	Johnson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris X.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwar C.,			
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			

Q8IMR8 drosophila
Q9VBE4 drosophila
Q8IMR9 drosophila
Q8GQ9 drosophila
Q8TNI4 methanosarc
Q9425 caenorhabdi
Q8C0E7 mus musculu
Q9V900 drosophila
Q9V780 drosophila
Q8C0E9 mus musculu
Q8F117 leptospira
Q9N18 dictyosteli
Q8F212 leptospira
Q8F2B3 leptospira
Q9D5G5 mus musculu
Q9D2F4 mus musculu
Q9P977 candida alb
Q9F411 candida alb
Q9V4C4 homo sapien
Q86QT0 drosophila
Q7Z5F7 homo sapien
Q7Z5F6 homo sapien
Q86Q87 drosophila
Q9N18 candida alb
Q96SW8 homo sapien
Q9NW37 homo sapien
Q9W2U2 drosophila
Q8F3F4 leptospira

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merklov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sidenkiam I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*;"
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003629; AAF52970.1; -
DR EMBL; AY058306; AAL13535.1; -
DR FlyBase; FBgn0010398; Lrr47.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR typ.
DR Pfam; PF00560; LRR; 4.
DR SMART; SM00369; LRR_TYP; 1.
SQ SEQUENCE 428 AA; 47961 MW; AED96CE6CA5B28D CRC64;

Query Match 16.3%; Score 284.5; DB 5; Length 428;
Best Local Similarity 28.7%; Pred. No. 3.8e-16;
Matches 96; Conservative 53; Mismatches 135; Indels 51; Gaps 10;

QY 15 LFTQAISSSLKGLFSAARMLAHRG-----CNVDTPVSTLTPTVKTSSEPFNKTKMVTSSKD 69
Db 91 LLIKCDPIQLKGLFQTLKLGMDGKDALNRLNNAATAIPQKAP-----QVRNVIKSE 146
QY 70 YPLSKNPPYSLEHLOTSYGGLVDMVMCLSKLRKLDLSHNHKKLPATIGLIHLQEL 129
Db 147 YPI-KGFPRTLKSITNNLSQVLKSFECTLRNLTKLDVSGNKLKIPSELGRL-PLTSL 204
QY 130 NLNDNHLESFS-----VALCHSTLQKSLRSLDLKNNKIKALPVQFCQLQELNKLDD 182
Db 205 HLGNNLLGTQNDWCWLRGTCLC-----QSLGELDLGNGLTYPPLVVKFESLVLNN 259
QY 183 NELIOPCKIGQLINLRFLSAARNKLPFLPSEFNNLSLEYDLFGNTFE-----QPKV 235
Db 260 NLLSLRPFPAIRMKALRKYVCSNELESLSAVEDLRIDLDVWGNCFKFNDAQAQMY 319
QY 236 LPVILKQAPLTLESARTILNRPYGSIIIPFHLCCQLDITAKICVCGRFCL----- 288
Db 320 LQKASNSQPLWLLGARAVDKYMLPLSAGSIIPAVLIDIREAPRCPCGELCYAQRKEDL 379
QY 289 -----NSFIQGTITNNL-----HVAHTVVLVDN 312
Db 380 FORVQPKFI---TVKNLYSREHQIYADVVLCD 411

RESULT 2
Q24386
ID Q24386 PRELIMINARY; PRT; 428 AA.
AC Q24386;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE LRR47 precursor.
GN LRR47 OR CG6098.

OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94289476; PubMed=8018718;
RA Ntwasa M., Buchanan S.G.S.C., Gay N.J.;
RT "Sequence and expression of LRR47, a novel embryonic leucine rich
RT repeat protein of *Drosophila*;"
RL Biochim. Biophys. Acta 1218:181-186(1994).
DR EMBL; X75760; CAA53387.1; -
DR FlyBase; FBgn0010398; Lrr47.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR typ.
DR Pfam; PF00560; LRR; 4.
DR SMART; SM00369; LRR_TYP; 1.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 428 AA; 47856 MW; CECEL29ALD7F8F7F CRC64;

Query Match 16.3%; Score 283.5; DB 5; Length 428;
Best Local Similarity 27.7%; Pred. No. 4.7e-16;
Matches 92; Conservative 55; Mismatches 140; Indels 45; Gaps 8;

QY 15 LFTQAISSSLKGLFSAARMLAHRG-----CNVDTPVSTLTPTVKTSSEPFNKTKMVTSSKD 69
Db 91 LLIKCDPIQLKGLFQTLKLGMDGKDALNRLNNAATAIPQKAP-----QVRNVIKSE 146
QY 70 YPLSKNPPYSLEHLOTSYGGLVDMVMCLSKLRKLDLSHNHKKLPATIGLIHLQEL 129
Db 147 YPI-KGFPRTLKSITNNLSQVLKSFECTLRNLTKLDVSGNKLKIPSELGRL-PLTSL 204
QY 130 NLNDNHLESFS-----VALCHSTLQKSLRSLDLKNNKIKALPVQFCQLQELNKLDD 182
Db 205 HLGNNLLGTQNDWCWLRGTCLC-----QSLGELDLGNGLTYPPLVVKFESLVLNN 259
QY 183 NELIOPCKIGQLINLRFLSAARNKLPFLPSEFNNLSLEYDLFGNTFE-----QPKV 235
Db 260 NLLSLRPFPAIRMKALRKYVCSNELESLSAVEDLRIDLDVWGNCFKFNDAQAQMY 319
QY 236 LPVILKQAPLTLESARTILNRPYGSIIIPFHLCCQLDITAKICVCGRFCL----- 290
Db 320 LQKASNSQPLWLLGARAVDKYMLPLSAGSIIPAVLIDIREAPRCPCGELCYAQRKEDL 379
QY 291 -----FIQGTITNNLHVAHTVVLVDN 312
Db 380 FORVQPKFISVKNLYSREHQIYADVVLCD 411

RESULT 3
Q8BX06
ID Q8BX06 PRELIMINARY; PRT; 366 AA.
AC Q8BX06;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical leucine-rich repeat.
GN C33001807RIK.
OS *Mus musculus* (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";

```

218 L-SLEVDLPGNTFEQKVLPIVK--LQAPLTLLLESSARTILHNRIPIYGSHLIIPHLQCD 274
266 LHLLELSLUGN-----PLVVRFRVDLYDPTLLELAARTIKIRINISTPVDLPGNLLRY 321
275 LDTAKIC---VCGRFCLNSFIQGTITTNKLHSAVHTVVLVDNLGCTEAPIISVFCSLGC 329
322 LGSASNCNPKCGGVYDCCVR-----QIKFVDFCGKYRULPMHYLCSPEC 367

RESULT 5
ID Q8F116 PRELIMINARY; PRT; 311 AA.
AC Q8F116;
DT 01-MAR-2003 (TrEMBLrel. 23, Created);
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update);
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update);
DE Leucine-rich repeat containing protein.
GN LA3323.
OS Leptosira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptosiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE011492; AAN50521.1; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_Typ.
DR Pfam; PF00560; LRR; 11.
DR PRINTS; PR00019; LEURICHRPT.
KW Complete proteome.
SQ SEQUENCE 311 AA; 35812 MW; 1D42A0EC3C40FCD4 CRC64;

```

```

Query Match          14.08; Score 244; DB 16; Length 311;
Best Local Similarity 32.6%; Pred. No. 7.9e-13;
Matches 74; Conservative 33; Mismatches 68; Indels 52; Gaps 6

QY 12 LQLLFQIAISSLKGPLSAMRLAHFGCNVDPVSTLPVKTSEFENPKTWITSKKDYP 71
Db 5 ITLNYLQKITIGL-----LFLMHSCE-----IQAEPEQGEYTDLTALQNP 48
QY 72 LSKNFPVSLBHLQNSYGLRVWRMMLCLKSLRKLDLSHNIKKLPATIGLIHQELNL 131
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
49 L-----KVRTLDLSANREKTLPKIEIGKLNKQLQELNL 79

QY 132 NDNHLESFVSVALCHSTLQKSLRSLDLSKNKTKALPVQFCQLQELKNLKLDDNELQFPECK 191
Db 80 NKNQLTTLPOEIQG---LKNLKSUNLSYNGIKTPEKTEKLOQSLGDDNNQLTLPQE 136

QY 192 IGOILNRLRFSAAFNKLPFLPSEPRLN-SLEYLDLFGNTFFQPKVLP 237
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
137 IGOILNQLSDLSSTNRLTTLPOEIGHLQNLQDLYGNSN---QLTILP 180

RESULT 6
Q8F3F9 PRELIMINARY; PRT; 498 AA.
ID Q8F3F9 AC Q8F3F9 AC
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative outer membrane protein.
GN LA2447.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011412; AAN49646.1; -.

```



```
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR typ.
DR InterPro; IPR000437; Prok_lipoprot_s.
DR Pfam; PF00560; LRR; 14.
DR PRINTS; PR00019; LEURICHRPT.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Complete proteome.
SQ SEQUENCE 498 AA; 57222 MW; E6D43494FD132FA73 CRC64;

Query Match 13.4%; Score 233.5; DB 16; Length 498;
Best Local Similarity 41.1%; Pred. No. 1.1e-11;
Matches 65; Conservative 29; Mismatches 59; Indels 5; Gaps 4;

QY 100 LKSLRKLDLSHNIKKLPATIGDLIHQELNLDNDHLESFVALCHSTLQKLSRLDLSK 159
DB 70 LKNQLNELNKNLLTVLPKEIQLENQLQELDRDQQLATFPVAVIVE--LQK-LESLLDSE 126
QY 160 NKIKALPVQFCOLQELKNLKLDDNELIOPCKIGQLINIRFLSAARNKLPFLPSFRNL- 218
DB 127 NRIILPLNEIGRLQNLQDLGLYKNKLTTPPKIGQLQNLQKLWLSNRUTALPKKEIGQLK 186
QY 219 SLEYLDLFGNTE-QPKVLPVILQAPLTLLESSARTI 255
DB 187 NLIQTLDQNOFTIILPKKEIQQLQNLQTLNLSNQLATL 224

RESULT 7
QY Q8F7S1 PRELIMINARY; PRT; 423 AA.
AC Q8F7S1;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Leucine-rich repeat containing protein.
GN LA0873.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011272; AAN48072.1; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR typ.
DR Pfam; PF00560; LRR; 13.
DR PRINTS; PR00019; LEURICHRPT.
KW Complete proteome.
SQ SEQUENCE 423 AA; 48695 MW; 91CB3DADB63C7DF0 CRC64;

Query Match 13.0%; Score 227; DB 16; Length 423;
Best Local Similarity 31.6%; Pred. No. 3.3e-11;
Matches 65; Conservative 44; Mismatches 89; Indels 8; Gaps 5;

QY 33 NYDTPVSTLTPVKTSEFENFKTMVITSKK-DYPLSKNFPYSLEHLQTSYCGLRVDMRM 97
DB 139 N-DNNKLTLPKEIQQLQNLQELSLNKLISLPTETIEQLKSLKLDLNLHNELTTSKEV 198
QY 98 LCLSKLRKLDLSHNIKKLPATIGDLIHQELNLDNDHLESFVALCHSTLQKLSRLDLS 157
DB 199 MLEETLENLDRSNKLTTPKEIQKLSKLVLMITGNQLTSLPKRI--EQQL-NLKTNL 255
QY 158 SKNKIKALPVQFCOLQELKNLKLDDNELIOPCKIGQLINIRFLSAARNKLPFLPSFRN 217
DB 256 GNRFOILSVEILDELKNLELNLYNQQLVEFPKVEGQLSKLYSLYHQNITTLPEVTQ 315
QY 218 L-SLEYLDLFGNTEQPKVLPVILQ 242
DB 316 LPDLQELHL-SGNKI---TILPKELIQ 338
```

RESULT 8

```
QY Q9N3F2 PRELIMINARY; PRT; 507 AA.
AC Q9N3F2;
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hypothetical protein.
GN Y54E10A.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Telodermata; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wollam A., Becker M., Graves T., Hawkins M.;
RT "The sequence of C. elegans cosmid Y54E10A."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC024810; AAF60763.1; -.
DR WormPep; Y54E10A.6; CE24450.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR typ.
DR Pfam; PF00560; LRR; 5.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 1.
KW Hypothetical protein.
SQ SEQUENCE 507 AA; 56414 MW; 98452F1FEC50591E CRC64;

Query Match 12.3%; Score 225.5; DB 5; Length 507;
Best Local Similarity 27.6%; Pred. No. 5.6e-11;
Matches 89; Conservative 40; Mismatches 114; Indels 79; Gaps 12;

QY 53 SEFENFKTMVITSKKDYPLSKNFPYSLEHLQTS-----YCGLRVDMRML-----C 99
DB 30 TEANLISQKRVFESMSQLNLLSLTGCSLHNLSSISCSNLMHLVLPKNDLKQLPDVFC 89
QY 100 LKSLRKLDLSHNIKKLPATIGDLIHQELNLDNDHLESFSA-----LCHST 147
DB 90 LPKLPKEDLSHNLDAIPASISKCNLESLLNNLNNELNESSFPDISNLSNHLIFDAHNT 149
QY 148 LQK-----SLRSLDLSKNIKALPVQFCOLQELKNLKLDDNELIOPCKIGQL 195
DB 150 ISKIPASLTSNLSAKLHTIILSHNIEVTPDSLSNKLQELKIDENKLDKVPVIAHL 209
QY 196 INLRFLSAARNKLPFLPSFRNLSLE-----YLDL-----FGNTEQPKV----- 235
DB 210 PKLVLDISKN--CFSESFPKLANDKRGKLNVAIALQKIGPKIGNSEEQKAPESGP 267
QY 236 -LPVILQAPLTLLESSARTILHN-----RIPYGSIIIPHLQDLDTAKICVGRCLNS 290
DB 268 DPV--PDAPLTV-----RTGIENLTVRRHPSVSEIRPVLVC-----CVINNVDLNG 312
QY 291 FIOGTTNLSHVAHTVLDVN 312
DB 313 GDSFKKFIALQTRKHASALCN 334

RESULT 9
QY Q8F3F6
```

```

ID Q8F3F6 PRELIMINARY; PRT; 633 AA.
AC Q8F3F6;
DT 01-MAR-2003 (TREMREL. 23, Created)
DT 01-MAR-2003 (TREMREL. 23, Last sequence update)
DT 01-OCT-2003 (TREMREL. 25, Last annotation update)
DE Putative outer membrane protein.
GN LA2450.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AE011413; AAN49649.1; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR Pfam; PF00560; LRR; 23.
DR PRINTS; PR00019; LEURICHRPT.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Complete proteome.
SQ SEQUENCE 633 AA; 72082 MW; 8C15836852DE2DDB CRC64;

Query Match 12.9%; Score 225; DB 16; Length 633;
Best Local Similarity 29.3%; Pred. No. 8e-11;
Matches 73; Conservative 43; Mismatches 89; Indels 50; Gaps 7;

QY 10 ILQLLFTQAISSSLKGFSLSAHLAHCNVDPVSTLTPVKTSFEN-----FKTKMV 63
DB 91 VIVELQKLESID-----LSENRL-----VMLPNEIGRLQLQELGLYKNKI 132
QY 64 ITSCKDYPKSNFPYSLEHLQ-SYGLVRVDMRLCKSLAKLDSLHNHKKLPATIGDL 123
DB 133 TFPKEIGQLQ-----NLQTLNQDNQATLTPVEIGQLQNLKLNKRLTVLPKEIGQL 187
QY 124 IHLQELNNDNHLFSFVALCH-STLQ-----KSLRSLDLKSNKIK 163
DB 188 QNLQTLNQDNQATLTPVEIGQLQNLQTLGSENQTLTPFKEIGQLQNLKLNKRLT 247
QY 164 ALPVQFQQLQELKMLDNDNLIQPPCKIGQLINRFLSAARKLPFPSPFNLSLEY 222
DB 248 ALPKEIGQLKLENLEUSENQTLTPFKEIGQLKLDLGLGRNQTLTPFKEIGQLKQLQ 307
QY 223 LDLFQNTPE 231
DB 308 LDLCYNQPK 316

RESULT 10
Q8F118 PRELIMINARY; PRT; 452 AA.
AC Q8F118;
DT 01-MAR-2003 (TREMREL. 23, Created)
DT 01-MAR-2003 (TREMREL. 23, Last sequence update)
DT 01-JUN-2003 (TREMREL. 24, Last annotation update)
DE Leucine-rich repeat containing protein.
GN LA3321.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB011492; AAN50519.1; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR.
DR Pfam; PF00560; LRR; 15.
DR PRINTS; PR00019; LEURICHRPT.
KW Complete proteome.

```

```

SQ SEQUENCE 452 AA; 52245 MW; 153B25C528FA21AA CRC64;

Query Match 12.9%; Score 224.5; DB 16; Length 452;
Best Local Similarity 30.1%; Pred. No. 5.9e-11;
Matches 74; Conservative 38; Mismatches 85; Indels 49; Gaps 8;

QY 46 TLTPVKTSFENFKT-----KMWITSKDYPLSKNFPYSLEHLQTSYCGLRVDMRLCL 100
DB 61 TTLPEIGQLQNLQRLDLSFNSLTLPKEIGQLRN-----LQELDLSEFNSLTLPKEVGQL 116
QY 101 KSLRKLDSLHNHKKLPATIGDLIHLQELNNDNHLSEFVALCHSTLCKSLR-----S 154
DB 117 ENLQRLDLHQNLATLPMEIGQLKMLQELCLNSKL-----TLPEIRQLRNLOE 167
QY 155 LDLSNNKIKALPVQFCQLQ-----ELKNLK-----LDDNELIQFPCK 191
DB 168 LDLSNNKLTLPKEIGQLQNLKTLNLTQTLTPKEIGQLQNLKTLNLDLQTLTPKE 227
QY 192 IGQLINRFLSAARKLPFPSPFNLSLEYDLFGNTFEO-PKVLPIVKLQAPITLLE 249
DB 228 IGELQNLLEILVLRNRLTALPEIKELQQLQNLQDLHQNLQTLTPKEIGQLQNLQRLDLHQ 287
QY 250 SSARTI 255
DB 288 NQITTL 293

RESULT 11
Q8F3F8 PRELIMINARY; PRT; 428 AA.
AC Q8F3F8;
DT 01-MAR-2003 (TREMREL. 23, Created)
DT 01-MAR-2003 (TREMREL. 23, Last sequence update)
DT 01-OCT-2003 (TREMREL. 25, Last annotation update)
DE Putative outer membrane protein.
GN LA2448.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AE011413; AAN49647.1; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR Pfam; PF00560; LRR; 15.
DR PRINTS; PR00019; LEURICHRPT.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Complete proteome.
SQ SEQUENCE 428 AA; 48954 MW; 0F7DD095D1BAC7C CRC64;

Query Match 12.8%; Score 222.5; DB 16; Length 428;
Best Local Similarity 31.8%; Pred. No. 8.2e-11;
Matches 74; Conservative 41; Mismatches 95; Indels 23; Gaps 8;

QY 39 NVDTPVSTLTPVKTSFEN-----FKTKWITSKDYPLSKNFPYSLEHLQTSYCG--- 89
DB 100 NLSSNQLTLPKEIGKLENLQRLDYDNLRT-----LPIEG---KLNQTLQTLSSNQ 151
QY 90 LVRVDMRLCKSLRKLDSLHNHKKLPATIGDLIHLQELNNDNHLSEFVALCHSTLQ 149
DB 152 LTTLPRESGKLENLQELNSDNQTLTPQELQQLQNLQTLNLSNQLTTFKEI---EGL 208
QY 150 KSLRSLDLKSNKIKALPVQFQQLQELKMLDNDNLIQPPCKIGQLINRFLSAARKLP 209
DB 209 KNLQTLNLSDNQTLTLPKEIGKLENLQELNSDNQTLTLPKEIGKLENLQELNSDNQTL 268
QY 210 FLPSFERNLSLEYDLFGNTFEO-PKVLPIVKLQAPITL-LESSARTILHNRI 260
DB 289 TLPIEIGKQLNHLNLSGNQTLTSLI-EIGKQLNQLNHLNLSNQLTILSKSI 320

```

```
RESULT 12
O61967 PRELIMINARY; PRT; 679 AA.
ID O61967 Q9NFN7;
AC 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 25, Last annotation update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein (LET-413 protein).
GN F26D11.11 OR LET-413.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2; PubMed=9851916;
RX MEDLINE=99069613;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Geisel C., Bradshaw H.;
RT "The sequence of C. elegans cosmid F26D11.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Legouis R., Sookharee S., Boshier J.M., Baillie D.L., Labouesse M.;
RT "LET-413 is a basolateral protein required for the assembly of
RT adherens junctions in C. elegans.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 2 PDZ/DHR DOMAINS.
DR EMBL; AF068716; AAC17752.2; -
DR EMBL; AJ276590; CAB91651.1; -
DR PIR; T33295; T33295.
DR WormPep; F26D11.11; CE29778.
DR GO; GO:007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR typ.
DR InterPro; IPR001478; PDZ.
DR Pfam; PF00560; PDZ; 11.
DR Pfam; PF00595; PDZ; 1.
DR SMART; SM00228; PDZ; 1.
DR PROSITE; PS0106; PDZ; 1.
KW Hypothetical protein.
SQ SEQUENCE 679 AA; 75335 MW; 608B8BDF5AC0F947 CRC64;
Query Match 12.5%; Score 218; DB 5; Length 679;
Best Local Similarity 39.4%; Pred. No. 3.5e-10;
Matches 61; Conservative 17; Mismatches 69; Indels 8; Gaps 3;
QY 80 LEHLQSYCVGLVRVDMRLCLKSLRKLDSLHNHKKLPATIGLIHLQELNNDNHLESF 139
DB 38 LEDLNLTNWNIKELDRHLSRLRLDLYSDNELAVLPAEIGNLTQILNLRN----- 92
QY 140 SVALCHSTLQ--KSLRSLDSLNKKIKALPVQFCOLQELNKLDDNELIQFCKIGQLIN 197
DB 93 SIAPLPDTPNQCKLTLTLNLSNPPFRLPETTCRCSITILSNLTSLLPSNIGSLTN 152
QY 198 LFLSLAARKLPFLPSEFRL--SLEYLDLFGNTFE 231
DB 153 LRVLEARDNLLATIPILSIVELAKLEELDIGQNELE 187

RESULT 13
O8F213 PRELIMINARY; PRT; 426 AA.
ID O8F213 Q8F213;
AC 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DB Leucine-rich repeat containing protein.
GN LA2964.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011460; AAN50163.1; -
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR typ.
DR Pfam; PF00560; LRR; 16.
DR PRINTS; PR00019; LEURICHRPT.
KW Complete proteome.
SQ SEQUENCE 426 AA; 49102 MW; 31342DCB537EEB70 CRC64;
Query Match 12.4%; Score 216; DB 16; Length 426;
Best Local Similarity 36.4%; Pred. No. 3e-10;
Matches 71; Conservative 34; Mismatches 66; Indels 24; Gaps 8;
QY 83 LQTSYCVGLVRVDMRLCLK--SLRKLDSLHNHKKLPATIGLIHLQELNNDNHLESF 139
DB 34 VQTYRNITEV-----LKNPCVNRILNLSGSKLITLPGSIGKLNQLNLDNQLIAL 87
QY 140 SVALCHSTLQKSLRSLDSLNKKIKALPVQFCOLQELNKLDDNELIQFCKIGQLINLR 199
DB 88 PKRI--GKLQ--NLQQLHLSKNQMLPFEISQGLNQLKLYENQLTAIPKEIGQLNQL 144
QY 200 FLSAARKLPFLPSEFRLNLS--LEVLDLFGNTFPQKVPVVKLQAPLTLESSARTILN 258
DB 145 ELNLNHLNQLATLPEDIEQLQRLQTLYLGNQFNS-----ILKEIGQLNQLNLSIG--LDEN 197
QY 259 RPYGSHIIPFHLCCQ 273
DB 198 QL-----NVLPKSIGQ 208

RESULT 14
O7ZSY7 PRELIMINARY; PRT; 289 AA.
ID O7ZSY7 Q7ZSY7;
AC 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TRISUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TRISUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041309; AAH41309.1; -
DR EMBL; BC044078; AAH44078.1; -
DR InterPro; IPR001611; LRR.
```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2004, 09:55:20 ; Search time 17 Seconds

(without alignments)
1035.277 Million cell updates/sec

Title: US-10-009-557-9

Perfect score: 1745

Sequence: 1 MYLSPVEFIILQLFIQAI.....PIISYFCSLGCYVNSDMLK 338

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1660	95.1	414	1 LLR1 HUMAN	Q96150 homo sapien
2	201.5	11.5	1256	1 LLR1 DROME	Q24020 drosophila
3	198	11.3	793	1 CDC1 HUMAN	Q9Y219 homo sapien
4	197.5	11.3	810	1 LRC8 HUMAN	Q81W66 homo sapien
5	194	11.1	810	1 LRC8 MOUSE	Q80W95 mus musculus
6	192.5	11.0	2026	1 CYAA YEAST	P08678 saccharomyc
7	183.5	10.5	276	1 RSU1 MOUSE	Q01730 mus musculus
8	183	10.5	1257	1 FLIH CAEL	P34268 caenorhabdi
9	180.5	10.3	276	1 RSU1 HUMAN	Q15404 homo sapien
10	178.5	10.2	582	1 SHO2 HUMAN	Q9UQ13 homo sapien
11	177.5	10.2	582	1 SHO2 MOUSE	O88520 mus musculus
12	175	10.0	1839	1 CYAA SACKL	P23466 saccharomyc
13	170	9.7	1039	1 YR71 CAEL	Q09564 caenorhabdi
14	169	9.7	371	1 LRC2 MOUSE	O8VDB8 mus musculus
15	168	9.6	371	1 LRC2 HUMAN	Q9BYS8 homo sapien
16	166.5	9.5	2493	1 CYAA USTWA	P49606 ustilago ma
17	165	9.5	1269	1 FLIH HUMAN	Q13045 homo sapien
18	164	9.4	2300	1 CYAA NEUCR	Q01631 neurospora
19	163.5	9.4	2145	1 CYAA PODAN	Q01513 podospora a
20	158	9.1	423	1 OMD MOUSE	O35103 mus musculus
21	158	9.1	581	1 LRG5 HUMAN	O8TF66 homo sapien
22	154	8.8	951	1 LGR4 HUMAN	Q9BXB1 homo sapien
23	154	8.8	1024	1 PCPC BALSO	O97B82 raistonia s
24	148.5	8.5	422	1 OMD BOVIN	O77742 bos taurus
25	148.5	8.5	560	1 GPV HUMAN	P40197 homo sapien
26	147	8.4	907	1 LGS5 MOUSE	Q921D4 mus musculus
27	146.5	8.4	421	1 OMD HUMAN	Q99383 homo sapien
28	146.5	8.4	662	1 GARP HUMAN	Q14392 homo sapien
29	146	8.4	828	1 LGS6 HUMAN	Q9BXB8 homo sapien
30	146	8.4	1054	1 R131 ARATH	Q91R14 arabidopsis
31	145	8.3	423	1 OMD RAT	Q921S7 rattus norv
32	144.5	8.3	1692	1 CYAA SCHPO	P14605 schizosach
33	144	8.3	343	1 LUM_CHICK	P51890 gallus gall

34	144	8.3	343	1 LUM_COTJA	Q9dec7 coturnix co
35	142.5	8.2	1192	1 EKS ARATH	Q91YN8 arabidopsis
36	141.5	8.1	578	1 LRJ5 RAT	O8r5m3 rattus norv
37	141	8.1	331	1 PLI8 AGKBL	O93233 agkistrodon
38	141	8.1	951	1 LGR4 RAT	Q9Z2H4 rattus norv
39	139	8.0	338	1 LUM_HUMAN	P51894 homo sapien
40	139	8.0	1207	1 BR11 LYCES	O8GUG5 lycopersico
41	139	8.0	1895	1 WR19 ARATH	Q9S267 arabidopsis
42	138.5	7.9	342	1 LUM_BOVIN	O05443 bos taurus
43	138.5	7.9	907	1 LGR5 HUMAN	O75473 homo sapien
44	138	7.9	782	1 CHAO TRICA	P82963 tribolium c
45	138	7.9	999	1 RLK5 ARATH	Z47735 arabidopsis

ALIGNMENTS

RESULT 1
LLR1 HUMAN
ID LLR1 HUMAN STANDARD; PRT; 414 AA.
AC Q96150; Q86SZ1; Q8N6H9;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE LRR-repeat protein 1 (LRR-1) (4-LBB-mediated signaling molecule);
DE (4-1BBLrr)
GN PPI15.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND CHARACTERIZATION.
RC TISSUE=T-cell lymphoma;
RX MEDLINE=21662677; PubMed=11804328;
RA Jang I.-K., Lee Z.-H., Kim H.-H., Hill J.M., Kim J.-D., Kwon B.S.;
RT "A novel leucine-rich repeat protein (LRR-1): potential involvement in
4-LBB-mediated signal transduction.";
Mol. Cells 12:304-312(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Cervical carcinoma;
RA Li W.B., Gruber C., Jessee J., Polayes D.;
RT "Full-length cDNA libraries and normalization.";
Submitted (Feb-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=B-cell;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
Klausner R.D., Collins F.S., Wagner J., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustun T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallat S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Maman A., Rodriguez S., Sanchez A.,
Whiting M., Maman A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley A.C., Touchman J.W., Green E.D., Dickinson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL FUNCTION: May negatively regulate the 4-LBB-mediated signaling
cascades which result in the activation of NK-kappaB and JNK1.
CC -!- SUBUNIT: Interacts with the cytoplasmic domain of TNFRSF9.
CC -!- ALTERNATIVE PRODUCTS:

```

CC CC Event=Alternative splicing; Named isoforms=2;
CC CC Name=1; Synonyms=LRR-1a;
CC CC IsoId=Q96L50-1; Sequence=Displayed;
CC CC Name=2; Synonyms=LRR-1b;
CC CC IsoId=Q96L50-2; Sequence=VSP_008363; VSP_008364;
CC CC -!- TISSUE SPECIFICITY: Ubiquitous. Maximal expression was seen in
CC CC the heart and skeletal muscle and minimal expression seen in
CC CC the kidney.
CC CC -!- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.
CC CC
CC CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC
CC CC -----
CC CC EMBL; AY052405; AAL11430.1; -
CC CC EMBL; BX248298; CAD62625.1; ALT_INIT.
CC CC EMBL; BC030142; RAH30142.1; -
CC CC Genew; HGNC:19742; PRL5.
CC CC InterPro; IPR001611; LRR.
CC CC Pfam; PF00560; LRR; 4.
CC CC PRINTS; PRO0019; LEURICHRPT.
CC CC Repeat; Leucine-rich repeat; Alternative splicing.
CC CC REPEAT 176 199
CC CC REPEAT 201 222
CC CC REPEAT 225 248
CC CC REPEAT 250 271
CC CC REPEAT 273 293
CC CC VARSPLIC 95 146
CC CC
CC CC TMVITSKXDY -> DSIMLSVHSIPSLPRFGYRKNLC:MK
CC CC IISLPHSRNYHESAPCCPHCGLSR (in isoform
CC CC 2).
CC CC
CC CC VARSPLIC 147 414
CC CC Missing (in isoform 2).
CC CC /FTID-VSP_008363.
CC CC /FTID-VSP_008364.
CC CC G -> V (IN REF. 1).
CC CC A -> S (IN REF. 1).
CC CC
CC CC CONFLICT 25 25
CC CC CONFLICT 28 28
CC CC CONFLICT 414 AA; 46722 MM; B19B17BD6CB33C4C CRC64;
CC CC
CC CC Query Match 95.1%; Score 1660; DB 1; Length 414;
CC CC Best Local Similarity 99.7%; Pred. No. 1.2e-120;
CC CC Matches 320; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC CC
CC CC QY 18 QAISSLKGFSLAMRLAHGCGNDTPVSTLTPVKTSEFENFKTKWITSKKYPLSKNFP 77
CC CC Db 94 KAISSLKGFSLAMRLAHGCGNDTPVSTLTPVKTSEFENFKTKWITSKKYPLSKNFP 153
CC CC
CC CC QY 78 YSLEHLQTSYCGLVVRDMRLCLSLRKLDSLNNHKKLPATIGDLIHQELMNDNHL 137
CC CC Db 154 YSLEHLQTSYCGLVVRDMRLCLSLRKLDSLNNHKKLPATIGDLIHQELMNDNHL 213
CC CC
CC CC QY 138 SFSVALCHSTLQKSLRSLDLKSKNKIKALPVQFQQLQELKNLKDDELIOFPCKIGOLIN 197
CC CC Db 214 SFSVALCHSTLQKSLRSLDLKSKNKIKALPVQFQQLQELKNLKDDELIOFPCKIGOLIN 273
CC CC
CC CC QY 198 LRFLSAARNKLPLPSEFNLSLEYLDLFGNTFQPKVLPVILQAPLTLLSSARTILH 257
CC CC Db 274 LRFLSAARNKLPLPSEFNLSLEYLDLFGNTFQPKVLPVILQAPLTLLSSARTILH 333
CC CC
CC CC QY 258 NRIPYGSIIHPHLCQDLDTAKICVGRFCNLSFIQGTITMNLHSAHTVVLVDNLGGTE 317
CC CC Db 334 NRIPYGSIIHPHLCQDLDTAKICVGRFCNLSFIQGTITMNLHSAHTVVLVDNLGGTE 393
CC CC
CC CC QY 318 APIISYFCSIGCYVNSDMLK 338
CC CC Db 394 APIISYFCSIGCYVNSDMLK 414
CC CC
CC CC RESULT 2
CC CC FLII_DROME

```

```

ID ID FLII DROME STANDARD; PRT; 1256 AA.
AC Q24020; Q24088; Q9VRH0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Flightless-I protein.
GN FLII OR CGI484.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Oregon-R; TISSUE=Embryo;
RX MEDLINE=94068608; PubMed=8248253;
RA Campbell H.D.; Schimanski T.; Claudianos C.; Ozsarac N.;
RA Kasprzak A.B.; Corseil C.N.; Young I.G.; de Couet H.G.; Miklos G.L.G.;
RT "The Drosophila melanogaster flightless-I gene involved in
RT gastrulation and muscle degeneration encodes gelolin-like and
RT leucine-rich repeat domains and is conserved in Caenorhabditis elegans
RT and humans."
RT Proc. Natl. Acad. Sci. U.S.A. 90:11386-11390 (1993).
RN [2]
SEQUENCE FROM N.A., AND VARIANT SER-601.
RC STRAIN=Canton-S;
RX MEDLINE=96129280; PubMed=8582612;
RA de Couet H.G.; Fong K.S.K.; Weeds A.G.; McLaughlin P.J.;
RA Miklos G.L.G.;
RT "Molecular and mutational analysis of a gelolin-family member encoded
RT by the flightless I gene of Drosophila melanogaster."
RL Genetics 141:1049-1059 (1995).
RN [3]
SEQUENCE FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=98188272; PubMed=9520435;
RA Maleszka R.; de Couet H.G.; Miklos G.L.G.;
RT "Data transferability from model organisms to human beings: insights
RT from the functional genomics of the flightless region of Drosophila."
RT Proc. Natl. Acad. Sci. U.S.A. 95:3731-3736 (1998).
RN [4]
SEQUENCE FROM N.A.
RC STRAIN=Berkely; TISSUE=Embryo;
RX MEDLINE=20136012; PubMed=10731138;
RA Rubin G.M.; Hong L.; Brokstein P.; Evans-Holm M.; Fzise E.;
RA Stapleton M.; Harvey D.A.;
RT "A Drosophila complementary DNA resource."
RL Science 287:2222-2224 (2000).
RN [5]
SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20136006; PubMed=10731132;
RA Adams M.D.; Celisner S.E.; Holt R.A.; Evans C.A.; Gocayne J.D.;
RA Amanatides P.G.; Scherer S.E.; Li P.W.; Hoskins R.A.; Galle R.F.;
RA Sutton G.G.; Lewis S.E.; Richards S.; Ashburner M.; Henderson S.N.;
RA Brannon R.C.; Rogers Y.-H.C.; Blazek R.G.; Zhang Q.; Chen L.X.;
RA Wan K.H.; Doyle C.; Baxter E.G.; Helt G.; Nelson C.R.; Miklos G.L.G.;
RA Abriel J.F.; Achayani A.; An H.-J.; Andrews-Pfannkuch C.; Baldwin D.;
RA Ballew R.M.; Basu A.; Bavendale J.; Bayraktaroglu L.; Beasley E.M.;
RA Beeson K.Y.; Benos P.V.; Berman B.P.; Bhandari D.; Bolshakov S.;
RA Borkova D.; Borchen M.R.; Bouck J.; Brokstein P.; Brottier P.;
RA Burtis K.C.; Busam D.A.; Butler H.; Cadieu E.; Center A.; Chandra I.;
RA Cherry J.M.; Cawley S.; Jahlke C.; Davenport L.B.; Davies P.;
RA de Pablos B.; Deicher A.; Deng Z.; Mays A.D.; Dew I.; Dietz S.M.;
RA Dodson K.; Doup L.E.; Downes M.; Dugan-Rocha S.; Dunkov B.C.; Duan P.;
RA Durbin K.J.; Evangelista C.C.; Ferraz C.; Ferreira S.; Fleischmann W.;
RA Fosler C.; Gabrielian A.E.; Garg N.S.; Gelbart W.M.; Glasser K.;
RA Glodek A.; Gong F.; Gorrell J.H.; Gu Z.; Guan P.; Harris M.;
RA Harris N.L.; Harvey D.A.; Heiman T.J.; Hernandez J.R.; Houck J.;
RA Hostin D.; Houston K.A.; Howland T.J.; Wei M.-H.; Ibegwam C.;
RA Jalali M.; Kalush F.; Karpen G.H.; Ke Z.; Kennison J.A.; Ketchum K.A.;
RA Kimmel B.E.; Kodira C.D.; Kraft C.; Kravitz S.; Kulp D.; Lai Z.;

```


CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; BC048152; AAR48152.1; -

DR InterPro; IPR001611; LRR.

DR InterPro; IPR003885; LRR_CYS.

DR InterPro; IPR003591; LRR_Typ.

DR 2fam; PR00560; LRR; 6.

DR PRINTS; PR00019; LEURICHRPT.

DR SMART; SM00366; LRR_PS; 8.

DR SMART; SM00365; LRR_SD22; 3.

DR SMART; SM00369; LRR_TVP; 8.

KW Repeat; Leucine-rich repeat; Signal.

FT SIGNAL 1 ? POTENTIAL.

FT CHAIN ? 810

FT REPEAT 38 62

FT REPEAT 422 445

FT REPEAT 492 515

FT REPEAT 518 542

FT REPEAT 543 565

FT REPEAT 590 612

FT REPEAT 613 637

FT REPEAT 639 660

FT REPEAT 661 684

FT REPEAT 706 706

FT REPEAT 707 730

FT REPEAT 732 752

FT REPEAT 753 776

FT REPEAT 778 801

FT CARBOHYD 66 66

FT CARBOHYD 83 83

FT CARBOHYD 191 191

FT CARBOHYD 526 526

FT CARBOHYD 554 554

FT CARBOHYD 592 592

FT CARBOHYD 660 660

FT CARBOHYD 696 696

FT CARBOHYD 755 755

SQ SEQUENCE 810 AA; 94119 MW; E3B125BC9977DBB0 CRC64;

Query Match 11.3%; Score 197.5; DB 1; Length 810;

Best Local Similarity 29.3%; Pred. No. 1.2e-07;

Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;

QY 57 NFKTKWITSKKDYPLSKNFPYSLSHLQTSYGLVVRDMRMLCKSLKRLDLS----- 109

Db 576 NEGTKLIVLN-----SUKKXVNLTELELRCDLERIPHSIFSLHNLQELDKNNLKI 629

QY 110 -----HNHIKKLPATIGDLIHQLQELMLNDHLESFVALCHSTLQKS 151

Db 630 EELISFQHLRLTCLKLNWNIAYIPIQIGNLTNLERLYLNENKIEKIPTQLFYC---RK 686

QY 152 LRSLDLNNKIKALPVQFCQLQELNKLDDNELIOPCKIQLINLRLSARNKLPPL 211

Db 687 LRYLDLSHNNLTFPADIGLGLQNLQNLAVTANRIEALPELPFCCKLRALHGNVQLSL 746

QY 212 PSEFENL-SLEYLDLFGNTFEQPKVLPVLIKQAPLTLLESSARTI---LHNRIIP 261

Db 747 PSRVGELNLQIELGRNLE---CLPVELGECPL-LKRSGLVVEEDLFTNLP 795

RESULT 5

ID LRCS MOUSE STANDARD; PRT; 810 AA.

AC Q80WG5;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Leucine-rich repeat-containing protein 8 precursor.

GN LRCS8.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_taxid=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altshuler R.D., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.M., Hong F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,

RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

CC -!- SIMILARITY: Contains 14 leucine-rich (LRR) repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; BC048152; AAR48152.1; -

DR InterPro; IPR001611; LRR.

DR InterPro; IPR003885; LRR_CYS.

DR InterPro; IPR003591; LRR_Typ.

DR 2fam; PR00560; LRR; 6.

DR PRINTS; PR00019; LEURICHRPT.

DR SMART; SM00366; LRR_PS; 8.

DR SMART; SM00365; LRR_SD22; 3.

DR SMART; SM00369; LRR_TVP; 8.

KW Repeat; Leucine-rich repeat; Signal.

FT SIGNAL 1 ? POTENTIAL.

FT CHAIN ? 810

FT REPEAT 38 62

FT REPEAT 422 445

FT REPEAT 492 515

FT REPEAT 518 542

FT REPEAT 543 565

FT REPEAT 590 612

FT REPEAT 613 637

FT REPEAT 639 660

FT REPEAT 661 684

FT REPEAT 706 706

FT REPEAT 707 730

FT REPEAT 732 752

FT REPEAT 753 776

FT REPEAT 778 801

FT CARBOHYD 66 66

FT CARBOHYD 83 83

FT CARBOHYD 191 191

FT CARBOHYD 526 526

FT CARBOHYD 554 554

FT CARBOHYD 592 592

FT CARBOHYD 660 660

FT CARBOHYD 696 696

FT CARBOHYD 755 755

SQ SEQUENCE 810 AA; 94119 MW; E3B125BC9977DBB0 CRC64;

Query Match 11.1%; Score 194; DB 1; Length 810;

Best Local Similarity 29.3%; Pred. No. 2.1e-07;

Matches 63; Conservative 30; Mismatches 84; Indels 38; Gaps 5;

QY 57 NFKTKWITSKKDYPLSKNFPYSLSHLQTSYGLVVRDMRMLCKSLKRLDLS----- 109

Db 576 NEGTKLIVLN-----SUKKXVNLTELELRCDLERIPHSIFSLHNLQELDKNNLKI 629

QY 110 -----HNHIKKLPATIGDLIHQLQELMLNDHLESFVALCHSTLQKS 151

Db 630 EELISFQHLRLTCLKLNWNIAYIPIQIGNLTNLERLYLNENKIEKIPTQLFYC---RK 686

QY 152 LRSLDLNNKIKALPVQFCQLQELNKLDDNELIOPCKIQLINLRLSARNKLPPL 211

Db 687 LRYLDLSHNNLTFPADIGLGLQNLQNLAVTANRIEALPELPFCCKLRALHGNVQLSL 746

QY 212 PSEFENL-SLEYLDLFGNTFEQPKVLPVLIKQAPL 245

Db 747 PSRVGELNLQIELGRNLE---CLPVELGECPL 778

RESULT 6

CYAA YEAST

ID CYAA YEAST STANDARD; PRT; 2026 AA.

AC P08678;

DT 01-JAN-1988 (Rel. 06, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Adenylate cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl cyclase).

GN CYR1 OR CDC35 OR HSR1 OR SRA4 OR YOL005W OR J1401.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86079531; PubMed=2934138;
RA Kataoka T., Broek D., Wigler M.;
RT "DNA sequence and characterization of the S. cerevisiae gene encoding
RL adenylylase cyclase";
RN Cell 43:493-505 (1985).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c / FY1679;
RA de Haan M., Smits P.H.M., Grivell L.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1042-2026 FROM N.A.
RX MEDLINE=88165073; PubMed=3327602;
RA Masson P., Lenz G., Jacquemin J.M., Danchin A.;
RT "Yeast adenylylase cyclase catalytic domain is carboxy terminal.";
RL Curr. Genet. 10:343-352 (1986).
RN [5]
RP MUTAGENESIS OF THR-1651.
RX MEDLINE=91122042; PubMed=1991451;
RA Feger G., de Vendittis E., Vitelli A., Masturzo P., Zahn R.,
RA Verrotti A.C., Kavounis C., Pal G.P., Fasano O.;
RT "Identification of regulatory residues of the yeast adenylyl
RL cyclase";
RN EMBO J. 10:349-359 (1991).
CC -!- FUNCTION: Plays essential roles in regulation of cellular
CC metabolism by catalyzing the synthesis of a second messenger,
CC cAMP.
CC -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
CC -!- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC -!- ENZYME REGULATION: The presence of GTP-bound Ras2 protein is
CC required in order to elicit a magnesium-dependent adenylyl cyclase
CC activity.
CC -!- SIMILARITY: Belongs to the adenylyl cyclase class-3 family.
CC -!- SIMILARITY: Contains 21 leucine-rich (LRR) repeats.
CC -!- SIMILARITY: Contains 1 PP2C-like domain.
CC -!- SIMILARITY: Contains 1 Ras-associating domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M12057; AAA34549.1; -;
CC EMBL; 249283; CAA89295.1; -;
CC EMBL; X87611; CAA60917.1; -;
CC EMBL; X03449; CAA27175.1; -;
CC PIR; S56776; OYBY
CC Germonline; 141621; -;
CC SGD; S0003542; CYR1.
CC GO; GO:0005886; C:plasma membrane; IDA.
CC GO; GO:0007126; P:meiosis; IMP.
CC GO; GO:0000074; P:regulation of cell cycle; IMP.
CC InterPro; IPR001054; G: cyclase.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR003591; LRR typ.
CC InterPro; IPR001932; PP2C-like.
CC InterPro; IPR000159; RA domain.
CC Pfam; PF00211; Guanylate_cyc; 1.
CC Pfam; PF00560; LRR; 15.

DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00044; CYC; 1.
DR SMART; SM00332; PP2C; 1.
DR SMART; SM00314; RA; 1.
DR PROSITE; PSS0125; GUANYLATE_CYCLASES_2; 1.
DR PROSITE; PSS0200; RA; 1.
KW Lyase; Repeat; Leucine-rich repeat; cAMP biosynthesis; Metal-binding;
KW Magnesium.
FT DOMAIN 676 755 RAS-ASSOCIATING.
FT REPEAT 815 838 LRR 1.
FT REPEAT 842 862 LRR 2.
FT REPEAT 863 885 LRR 3.
FT REPEAT 886 908 LRR 4.
FT REPEAT 910 931 LRR 5.
FT REPEAT 932 955 LRR 6.
FT REPEAT 957 976 LRR 7.
FT REPEAT 977 999 LRR 8.
FT REPEAT 1001 1016 LRR 9.
FT REPEAT 1017 1040 LRR 10.
FT REPEAT 1042 1062 LRR 11.
FT REPEAT 1063 1086 LRR 12.
FT REPEAT 1088 1109 LRR 13.
FT REPEAT 1110 1132 LRR 14.
FT REPEAT 1134 1156 LRR 15.
FT REPEAT 1188 1209 LRR 16.
FT REPEAT 1210 1232 LRR 17.
FT REPEAT 1233 1256 LRR 18.
FT REPEAT 1258 1280 LRR 19.
FT REPEAT 1285 1308 LRR 20.
FT REPEAT 1319 1344 LRR 21.
FT DOMAIN 1369 1625 PP2C-LIKE.
FT DOMAIN 1626 2026 CATALYTIC.
FT METAL 1673 1673 MAGNESIUM (BY SIMILARITY).
FT METAL 1716 1716 MAGNESIUM (BY SIMILARITY).
FT MUTAGEN 1651 1651 T->I: WEAK RAS-INDEPENDENT ACTIVITY.
FT MUTAGEN 1651 1651 T->D: ATTENUATION OF THE RESPONSE TO
FT RAS PROTEINS.
FT CONFLICT 262 262 S -> L (IN REF. 1).
FT CONFLICT 548 548 S -> L (IN REF. 1).
FT CONFLICT 592 592 D -> H (IN REF. 1).
FT CONFLICT 709 709 R -> I (IN REF. 1).
FT CONFLICT 962 962 L -> P (IN REF. 1).
FT CONFLICT 1388 1388 L -> S (IN REF. 1).
FT CONFLICT 1427 1427 E -> D (IN REF. 3).
FT CONFLICT 1461 1461 A -> T (IN REF. 3).
FT CONFLICT 1566 1566 A -> S (IN REF. 1).
FT CONFLICT 1735 1735 V -> G (IN REF. 1).
FT CONFLICT 1956 1956 I -> V (IN REF. 3).
FT CONFLICT 1996 1996 C -> F (IN REF. 1).
FT CONFLICT 2009 2026 NVYDELLQWVKNKADLST -> MLLTNFYKMLETQRIYQL
FT CONFLICT 2026 2026 FCS (IN REF. 3).
SQ SEQUENCE 2026 AA; 227832 MW; BEAB7419AB6989D0 CRC64;
Query Match 11.0%; Score 192.5; DB 1; Length 2026;
Best local Similarity 25.8%; Pred. No. 8.4e-07;
Matches 84; Conservative 52; Mismatches 125; Indels 63; Gaps 13;
QY 2 YVLSPVETIIQLLFTQALSSSLKGLSAMLAHRCNVDPVSTLTVEVKTSEFENFKTK 61
DB 863 YKLVSLE---LQNFIRKVPNSIMK-LSNLTINLQCN---ELESF-PAGFVELKQLQL 914
QY 62 MVTSK-KDYPLSKNPPYSLEHLQTSYGLVRVDMMLCLSKLRKLDLGHNHKKLPAT 120
DB 915 DLSNNKFMHYPEVINYCTNLQIDLQSLQSLQSTKYLVKLVKLNLSHNKLN---FI 970
QY 121 GDL-----IHLQELNLDNHLESFVAL----- 143
DB 971 GDLSEMTDLRTNLNRYNRISSTKTNASNLQNLFLTDNRISNEEDTLPKLRALEIOENPIT 1030
QY 144 ---CHSTLQKSLRSLDLKSKKIKALPVQ-FCOLOELKLNKLDNDELIOFPCKIGQLINLR 199
DB 1031 SISFKDFYFKNWTSLTNKKAQLSSIPGELLTKLSFLEKLELNQNLTRLPQELSKTKLV 1090

QY 230 FLSSARNKLPFLPSPERNL-SLEYDLFGNTF-----EQPKVLPVVKLQAPLTL-ESSAR 253
 DB 1091 FLSSARNKLEYIPPELSQKSLRTLDLHNHNRDQFVDMENLELTSLTSSNAFGNSSLE 1150
 QY 254 TLHNRIPIVGSHT---IPHLQCD 274
 DB 1151 NSFYHNSYSGSKLSKSLMFFIAAD 1174

RESULT 7
 RSUL_MOUSE
 ID RSUL_MOUSE STANDARD; PRT; 276 AA.
 AC Q01730;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ras suppressor protein 1 (Rsu-1) (RSP-1).
 GN RSU1 OR RSP1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIH Swiss; TISSUE=Fibroblast;
 RX MEDLINE=92375042; PubMed=1508180;
 RA Cutler M.L., Bassin R.H., Zannoni L., Talbot N.;
 RT "Isolation of rsp-1, a novel cDNA capable of suppressing v-Ras
 transformation.";
 RL Mol. Cell. Biol. 12:3750-3756 (1992).
 CC -!- FUNCTION: Potentially plays a role in the Ras signal transduction
 pathway. Capable of suppressing v-Ras transformation in vitro.
 CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DB EMBL: X63039; GAA44765.1; -
 DR PIR: S25770; S25770.
 DR MGD: MGI:103040; Rsul.
 DR GO: GO:0037265; P-RAS protein signal transduction; IMP.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR003591; LRR_TYP.
 DR Pfam: PF00560; LRR; 7
 DR SMART: SM00364; LRR_BAC; 5.
 DR SMART: SM00366; LRR_PS; 6.
 DR SMART: SM00369; LRR_TYP; 6.
 DR Leucine-rich repeat; Repeat.
 DR INIT MET 0 BY SIMILARITY.
 FT REPEAT 38 61 LRR 1.
 FT REPEAT 63 84 LRR 2.
 FT REPEAT 85 107 LRR 3.
 FT REPEAT 109 132 LRR 4.
 FT REPEAT 133 155 LRR 5.
 FT REPEAT 157 178 LRR 6.
 FT REPEAT 179 201 LRR 7.
 SQ SEQUENCE 276 AA; 31419 MW; 245FE4AA27AF16DE CRC64;

Query Match 10.5%; Score 183.5; DB 1; Length 276;
 Best Local Similarity 31.2%; Pred. No. 3.8e-07;
 Matches 60; Conservative 26; Mismatches 81; Indels 25; Gaps 3;
 QY 72 LSKNFPYSLEH---LQTSYCYGLVVDVMDMLCKSLRKLDSLHNHKKUPATIGDLIHQOE 128
 DB 30 LDVNGDLSLAHTQTMLSHNKLTTVPNVVAELKKNLELVNFFNNQIELEPTQISSIQKLKH 89
 QY 129 LNLNDNHEGFS-----VALCHSTLQK-----SLRSLDSKKNKALP 166

DB 90 INLGMNRLNTLPRGFGSSRLLEVLTYNMLNHSHPGNFFYLTTLRALYLSNDFFILP 149
 QY 167 VQFCQELKKNLKDNLQIQPKIGQLINRFLSAARKLPFLPSEFRNLSLEYDLDF 226
 DB 150 PDIGKLTQKQLSLRNDLISLPKEIGELTQKELHIQGNRLTVLPPELGNLDLTGQKQV 209
 QY 227 GNTFEQPKVLPV 238
 DB 210 FKAENNEWVTEI 221

RESULT 8
 FLIH_CABEL
 ID FLIH_CABEL STANDARD; PRT; 1257 AA.
 AC P34268; P90739; Q1731;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Flightless-I protein homolog.
 GN B0523.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=94068608; PubMed=8248259;
 RX Campbell H.D., Schimansky T., Claudianos C., Ozsarac N.,
 RA Kasprzak A.B., Cotwell J.N., Young I.G., de Couet H.G., Miklos G.L.G.;
 RT "The Drosophila melanogaster flightless-I gene involved in
 gastrulation and muscle degeneration encodes gelsolin-like and
 leucine-rich repeat domains and is conserved in Caenorhabditis
 elegans and humans.";
 RT Proc. Natl. Acad. Sci. U.S.A. 90:11386-11390 (1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Altschough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S.D., Z., Durbin R., Favell A., Fraser A.,
 RA Fulston L., Gardner A., Green P., Hawkins J., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38 (1994).
 RN [3]
 RP REVISIONS.
 RA Waterston R.;
 RL Submitted (JUN-1996) to Swiss-Prot.
 CC -!- FUNCTION: May play a key role in embryonic cellularization by
 interacting with both the cytoskeleton and other cellular
 components (By similarity).
 CC -!- SIMILARITY: Belongs to the villin/gelsolin family.
 CC -!- SIMILARITY: Contains 16 leucine-rich (LRR) repeats.
 CC -!- SIMILARITY: Contains 4 gelsolin-like repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DB EMBL: U01183; AAC03567.1; -
 DB EMBL: L07143; AAB37088.1; -

RI	LOE
RL	DNA
RN	[4]

```
Query Match      10.2%; Score 178.5; DB 1; Length 582;
Best Local Similarity 30.8%; Pred. No. 2.3e-06;
Matches 57; Conservative 26; Mismatches 63; Indels 39; Gaps 6;

QY 102 SRRKLDLSNHKKLPATIGDLHLHQLNLDNHLSPFVALCHSTLQKSLRSLDLSKNK 161
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 403 SWEVLNLTATNQTLPEDVSGVLSLEVLILSNLLKLPGL--GNLRK--LRDLLEENK 459
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 162 IKALPVQFCOLOELKLNKLDNDELIOFPCKIGOLINLRFLSAARNKLPFLPSEFRNL-SL 220
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 460 LESLPNEIAYLKDLQKLVLTNNQTLTPRGIGHTLNTLHGLGENLLTHLPEEIGTLENL 519
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 221 EYLDLFGNTPEQKVPVKLPKQAPLFLLESSARTILHNRIPIYGSHPH--LCQDLDTA 278
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 520 EELYNDN-----PNT-----HSLPFELALCSKLSIM 546
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 279 KICVC 283
   |
DB 547 SIENC 551

RESULT 11
SHO2 MOUSE
ID SHO2 MOUSE STANDARD; PRT; 582 AA.
AC O88520;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Leucine-rich repeat protein SHOC-2 (Ras-binding protein Sur-8).
GN SHOC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98337190; PubMed=9674433;
RA Sieburth D.S., Sun Q., Han M.;
RT "SUR-8, a conserved Ras-binding protein with leucine-rich repeats,
RL Cell 94:119-130(1998).
CC -! SUBUNIT: Specifically binds K-Ras and N-Ras but not H-Ras in
CC vitro.
CC -! SIMILARITY: Contains 20 leucine-rich (LRR) repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF068921; AAC40175.1; -
CC MGD; MG1:1927197; Shoc2.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR003591; LRR_TYP.
CC Pfam; PF00560; LRR; 16.
CC PRINTS; PR00019; LEURICRPT.
CC SMART; SM00369; LRR_TYF; 1.
CC Repeat; Leucine-rich repeat.
FT REPEAT 99 122 LRR 1.
FT REPEAT 123 145 LRR 2.
FT REPEAT 146 168 LRR 3.
FT REPEAT 169 191 LRR 4.
FT REPEAT 193 214 LRR 5.
FT REPEAT 215 237 LRR 6.
FT REPEAT 239 260 LRR 7.
FT REPEAT 262 283 LRR 8.
FT REPEAT 284 306 LRR 9.
FT REPEAT 307 329 LRR 10.
FT REPEAT 331 353 LRR 11.
FT REPEAT 354 377 LRR 12.

Query Match      10.2%; Score 177.5; DB 1; Length 582;
Best Local Similarity 30.8%; Pred. No. 2.7e-06;
Matches 57; Conservative 26; Mismatches 63; Indels 39; Gaps 6;

QY 102 SRRKLDLSNHKKLPATIGDLHLHQLNLDNHLSPFVALCHSTLQKSLRSLDLSKNK 161
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 403 SWEVLNLTATNQTLPEDVSGVLSLEVLILSNLLKLPGL--GNLRK--LRDLLEENK 459
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 162 IKALPVQFCOLOELKLNKLDNDELIOFPCKIGOLINLRFLSAARNKLPFLPSEFRNL-SL 220
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 460 LESLPNEIAYLKDLQKLVLTNNQTLTPRGIGHTLNTLHGLGENLLTHLPEEIGTLENL 519
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 221 EYLDLFGNTPEQKVPVKLPKQAPLFLLESSARTILHNRIPIYGSHPH--LCQDLDTA 278
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 520 EELYNDN-----PNT-----HSLPFELALCSKLSIM 546
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 279 KICVC 283
   |
DB 547 SIENC 551

RESULT 12
CYAA SACKL
ID CYAA SACKL STANDARD; PRT; 1839 AA.
AC P23466;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Adenylate cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl-
DE cyclase).
GN CYR1
OS Saccharomyces kluyveri (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4934;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91323718; PubMed=1864503;
RA Young D., O'Neill K., Broek D., Wigler M.;
RT "The adenylyl cyclase-encoding gene from Saccharomyces kluyveri."
RL Gene 102:129-132(1991).
CC -! FUNCTION: Plays essential roles in regulation of cellular
CC metabolism by catalyzing the synthesis of a second messenger,
CC CAMP.
CC -! CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
CC -! COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC -! SIMILARITY: Belongs to the adenylyl cyclase class-3 family.
CC -! SIMILARITY: Contains 21 leucine-rich (LRR) repeats.
CC -! SIMILARITY: Contains 1 PP2C-like domain.
CC -! SIMILARITY: Contains 1 Ras-associating domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X56042; CAA39513.1; -
CC InterPro; IPR001054; G_cyclase.
CC InterPro; IPR001611; LRR.
```

```

DR InterPro; IPR003591; LRR typ.
DR InterPro; IPR001932; PP2C-like.
DR InterPro; IPR000159; RA domain.
DR Pfam; PF00211; Guanylate_cyc; 1.
DR Pfam; PF00560; LRR; 15.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00044; CYCC; 1.
DR SMART; SM00369; LRR_TYP; 2.
DR SMART; SM00332; PP2C; 1.
DR SMART; SM00314; RA; 1.
DR PROSITE; PS0125; GUANYLATE_CYCLASES_2; 1.
DR PROSITE; PS0200; RA; 1.
DR Lyase; Repeat; Leucine-rich repeat; cAMP biosynthesis; Metal-binding;
KW Magnesium.
FT DOMAIN 494 574 RAS-ASSOCIATING.
FT REPEAT 632 655 LRR 1.
FT REPEAT 659 679 LRR 2.
FT REPEAT 680 702 LRR 3.
FT REPEAT 703 724 LRR 4.
FT REPEAT 726 748 LRR 5.
FT REPEAT 749 771 LRR 6.
FT REPEAT 773 793 LRR 7.
FT REPEAT 794 818 LRR 8.
FT REPEAT 820 833 LRR 9.
FT REPEAT 834 857 LRR 10.
FT REPEAT 859 879 LRR 11.
FT REPEAT 880 903 LRR 12.
FT REPEAT 905 926 LRR 13.
FT REPEAT 927 949 LRR 14.
FT REPEAT 951 974 LRR 15.
FT REPEAT 1004 1025 LRR 16.
FT REPEAT 1026 1048 LRR 17.
FT REPEAT 1050 1072 LRR 18.
FT REPEAT 1074 1096 LRR 19.
FT REPEAT 1101 1124 LRR 20.
FT REPEAT 1135 1160 LRR 21.
FT DOMAIN 1185 1440 PP2C-LIKE.
FT DOMAIN 1441 1839 CATALYTIC.
FT METAL 1488 1488 MAGNESIUM (BY SIMILARITY).
FT METAL 1531 1531 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 1839 AA; 206895 MW; 86A69BCB1F273CB CRC64;

Query Match 10.0%; Score 175; DB 1; Length 1839;
Best Local Similarity 26.3%; Pred. No. 1.7e-05;
Matches 69; Conservative 49; Mismatches 88; Indels 56; Gaps 11;

QY 12 LQLLFIQAISSSLKGF-LSAMELAHRCNVDPVSTLTPTVKTSPEFNFKTKWITSKK-- 68
Db LERNFIKKVPDSI--FKLNNLTIVNLCN--NLRELPP-GFSLKXNLQL-LDISNKFV 739
QY 69 DYPLSKNFPYSLEHLOTSYCGLVVRVDMRMLCLKSLRKLDSLHNHKKLPATIGDLI--- 124
Db DYPLSKNFPYSLEHLOTSYCGLVVRVDMRMLCLKSLRKLDSLHNHKKLPATIGDLI--- 124
QY 740 NYPEVINSCTNLQIDLSYNKTHSLPVSINQLVKLAKNLFNNRL----TSVGDLSQMK 795
Db NYPEVINSCTNLQIDLSYNKTHSLPVSINQLVKLAKNLFNNRL----TSVGDLSQMK 795
QY 125 -----HLOELNNDNHLSEF-----SVALCHSTL 148
Db LRTNLCRNVTSTCHAPFNQNLNLTNRTSTFDDTLRLTLELQQNPISWVCGGNY 855
QY 149 QKSLESLDSKNNKIKALPVQ-FCQQLKXNLKLDNDELITQFCKTGQNLNLFLSA 207
Db QKSLESLDSKNNKIKALPVQ-FCQQLKXNLKLDNDELITQFCKTGQNLNLFLSA 207
QY 856 MANWTSLSLNKAKLSFSAELLSKLPRLKLELNENNLTQLPPEINKLRLYLVSARNK 915
Db MANWTSLSLNKAKLSFSAELLSKLPRLKLELNENNLTQLPPEINKLRLYLVSARNK 915
QY 208 LPFLSEFRNL-SLEYLDLFGN 228
Db LPFLSEFRNL-SLEYLDLFGN 228
QY 916 LESIPDEISDRLSKSLDLSN 937
Db LESIPDEISDRLSKSLDLSN 937

RESULT 13
YR71_CABEL
ID YR71_CABEL STANDARD; PRT; 1039 AA.
AC Q09564;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)

```

```

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein F43C1.1 in chromosome III.
GN F43C1.1
OS Caenorhabditis elegans.
OC Rukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Jassal B.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -| SIMILARITY: Contains 18 leucine-rich (LRR) repeats.
CC -| SIMILARITY: Contains 1 PP2C-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; Z46937; CAA87056.1; -.
DR PIR; T22117; T22117.
DR WormPep; F43C1.1; CE01582.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR typ.
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00560; LRR; 13.
DR Pfam; PF00481; PP2C; 1.
DR PRINTS; PR00319; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 1.
DR SMART; SM00332; PP2C; 1.
KW Hypothetical protein; Leucine-rich repeat; Repeat.
FT REPEAT 137 160 LRR 1.
FT REPEAT 161 185 LRR 2.
FT REPEAT 199 222 LRR 3.
FT REPEAT 223 245 LRR 4.
FT REPEAT 246 268 LRR 5.
FT REPEAT 269 291 LRR 6.
FT REPEAT 293 316 LRR 7.
FT REPEAT 332 358 LRR 8.
FT REPEAT 364 389 LRR 9.
FT REPEAT 396 419 LRR 10.
FT REPEAT 420 442 LRR 11.
FT REPEAT 444 464 LRR 12.
FT REPEAT 465 487 LRR 13.
FT REPEAT 489 514 LRR 14.
FT REPEAT 536 559 LRR 15.
FT REPEAT 561 585 LRR 16.
FT REPEAT 587 605 LRR 17.
FT REPEAT 606 627 LRR 18.
FT DOMAIN 669 903 PP2C-LIKE.
SQ SEQUENCE 1039 AA; 118182 MW; 0C6D1AFC79A0D32A CRC64;

Query Match 9.7%; Score 170; DB 1; Length 1039;
Best Local Similarity 25.3%; Pred. No. 2e-05;
Matches 77; Conservative 52; Mismatches 121; Indels 54; Gaps 11;

QY 24 LKGFISAMLAHRCNVDPVSTLTPTVKTSPEFNFKTKWITSKKDYPLSKNFPYSLEHL 83
Db LEAWLTRAQQVEKSNHVDASDEQLTLEPQILNNEARVQILNLRNLSLRP---PTKVV 183
QY 84 QTSYCGLVVRVDMRMLCLKSLRKLDSLHNHKKLPATIGDLIHLQELNNDNHLSEFVAL 143
Db QTSYCGLVVRVDMRMLCLKSLRKLDSLHNHKKLPATIGDLIHLQELNNDNHLSEFVAL 143
QY 184 QKSMAPLGYID-----DLYRVH-----SLQVDSLANSQILSPFIQL 219
Db QKSMAPLGYID-----DLYRVH-----SLQVDSLANSQILSPFIQL 219
QY 144 CHSTLQKSLRSLDSKNNKIKALPVQFCQQLKXNLKLDNDELITQFCKTGQNLNLFLSA 203
Db CHSTLQKSLRSLDSKNNKIKALPVQFCQQLKXNLKLDNDELITQFCKTGQNLNLFLSA 203
QY 220 ---TLSSHRLQNLSSNTYSSVPSSECSNMRRQLQYLNLSNNQLDT-PSDISELQNSLDI 276
Db ---TLSSHRLQNLSSNTYSSVPSSECSNMRRQLQYLNLSNNQLDT-PSDISELQNSLDI 276
QY 284 ARNKK-FFLPSEFRNLSLEYLDLFGNTTFQ-PKV--LPVTKLQAP-TLLESSAR----TIL 256
Db ARNKK-FFLPSEFRNLSLEYLDLFGNTTFQ-PKV--LPVTKLQAP-TLLESSAR----TIL 256

```

```

Db      277  SENQSQIPPCPLPHLTLEMLAGNIEKIDRGEMQIQKIDLRNVLDTSFRLDIENT 336
QY      257  HNRIPYGVSHIIPHLQODLTKICVGRFCLNSFQIGITIMNLH-----SVAHTVV---L 309
Db      337  HLDLRNLSMISTVHL-TNLRFLVKVHCERLQSS-----IHLGSESLTEVYADHNL 386
QY      310  VDNL 313
Db      387  LDSL 390

RESULT 14
LRC2 MOUSE
ID      LRC2 MOUSE STANDARD; PRT; 371 AA.
AC      Q8VDB8; Q9CX04;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Leucine-rich repeat-containing protein 2.
GN      LRR2.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=BALB/C;
RA      Kiss H., Darai E., Kiss C., Kost-Alimova M., Klein G., Dumanski J.P.,
RA      Imreh S.;
RT      "Comparative sequence analysis of the common eliminated region 1
RT      (C3CER1) on human chromosome 3p21.3 with the orthologous regions on
RT      mouse chromosome 9p."
RL      Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
[2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CS7BL/6J; TISSUE=Embryonic stem cells;
RX      MEDLINE=21085660; PubMed=11217851;
RA      Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA      Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA      Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA      Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA      Kadota K., Matsuda H.A., Asburner M., Batalov S., Casavant T.,
RA      Fleischmann W., Gaasterland T., Glessi C., King B., Kochiwa H.,
RA      Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA      Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA      Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh S.,
RA      Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA      Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA      Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA      Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA      Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA      Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA      Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA      Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseki S.,
RA      Hayashizaki Y.;
RT      "Functional annotation of a full-length mouse cDNA collection.";
RL      Nature 409:685-690 (2001).
CC      - SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; AJ428068; CAD20990.1; -
CC      EMBL; AK010252; BAB26797.1; -
CC      MGD; MGI:1921499; Lrrc2.
CC      InterPro; IPR001611; LRR.
CC      InterPro; IPR003591; LRR_typ.

```

```

DR      PFam; PFO0560; LRR; 3.
DR      SMART; SM00369; LRR_Typ; 1.
KW      Repeat; Leucine-rich repeat.
FT      REPEAT 97 120 LRR 1.
FT      REPEAT 121 143 LRR 2.
FT      REPEAT 144 166 LRR 3.
FT      REPEAT 167 189 LRR 4.
FT      REPEAT 191 213 LRR 5.
FT      REPEAT 214 236 LRR 6.
FT      REPEAT 237 259 LRR 7.
FT      REPEAT 261 282 LRR 8.
FT      REPEAT 283 305 LRR 9.
FT      CONFLICT 52 52 Q -> E (IN REF. 2).
SQ      SEQUENCE 371 AA; 42955 MW; 55E96FF2D60E6A23 CRC64;

Query Match 9.7%; Score 169; DB 1; Length 371;
Best Local Similarity 25.9%; Pred. No. 7.1e-06;
Matches 70; Conservative 37; Mismatches 91; Indels 72; Gaps 10;

QY      40  VDPVSTLTPVKTSEFENFKMTKVTSSKKDYP-----LS-----KNFPYSLS---HLQTSY- 87
Db      69  VDTSTLLEKIERNSYSRQSRVQKDRKRSPPVFELSGTQWKELPDSLKEQTHLKEWYI 128
QY      88  -CGLVWDVWMLCL-KSLRKLDLSHNIKKLPATIGDLHLQELNLDNHNLESFVAL-- 143
Db      129  HSTLIQIIPTYIELFQAMKILDLPKQKITCLPAEIGRLKMLKELNVSFNHLKSIPELGD 188
QY      144  -----CHSTLQ-----KSLRSLDLSKNKIKALPVQFCQLQELKMLKLDNNE 184
Db      189  CEHLERLDCGNLDLMDLPFELSNLQVTFVDISANKFSVPICVLRMCRLQWLDISSNN 248
QY      185  LIQFPCKIGQLNLRFLSARKNLPFLPSEFRNLSLEYLDLFGNTFEPQKVPVLKQAP 244
Db      249  LSDLPQDIDRLBELQGFLYKKNKLYLPQAMLNK-----K 284
QY      245  LTLLESSARTILHNRIYPYGSHI--IPFHL 272
Db      285  LTLVVS-----GDHLVEVPTALC 303

RESULT 15
LRC2_HUMAN
ID      LRC2_HUMAN STANDARD; PRT; 371 AA.
AC      Q9BY58; Q96LIT5;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Leucine-rich repeat-containing protein 2.
GN      LRR2.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=21906202; PubMed=11896456;
RA      Kiss H., Yang Y., Kiss C., Andersson K., Klein G., Imreh S.,
RA      Dumanski J.P.;
RT      "The transcriptional map of the common eliminated region 1 (C3CER1) in
RT      3p21.3.";
RL      Eur. J. Hum. Genet. 10:52-61 (2002).
[2]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Cerebellum;
RA      Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Suzuki Y.,
RA      Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,
RA      Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,
RA      Kawakami B., Nagai K., Isogai T., Sugano S.;
RT      "NEDO human cDNA sequencing project.";
RL      Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
[3]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Colon, Kidney, and Stomach;

```



```

QY 241 LQAPITLLESSARTILNRPYQSHI--PFIHLC 272
      ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 284 ---KZILLVVS-----GDHLWELTALC 303

Search completed: June 15, 2004, 10:01:48
Job time : 18 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2004, 09:59:10 / Search time 21 seconds
(without alignments)
1548.226 Million cell updates/sec

Title: US-10-009-557-9

Perfect score: 1745
Sequence: 1 MYVLSPVEFIILQLFIQAI.....PIISVFCSLGCVNSSDMLX 338

Scoring table: BLOSUM62
Gapop 10.0, Gapext 3.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR 78.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	283.5	16.2	428	2 S45361	LR47 protein - fr
2	218	12.5	717	2 T33295	hypothetical prote
3	208.5	11.9	1495	2 T31434	densin-180 - rat
4	201.5	11.5	1256	2 S60461	gene flightless-I
5	196	11.2	492	2 P86283	hypothetical prote
6	192.5	11.0	886	2 T40734	probable adenylate
7	192.5	11.0	2026	1 OYBY	adenylate cyclase
8	191	10.9	380	2 T01281	probable leucine-r
9	189.5	10.9	526	2 C84552	hypothetical prote
10	189	10.8	1257	2 A88536	protein B0523.5 [i
11	183.5	10.5	277	2 S25770	RSP-1 protein - mo
12	181.5	10.4	549	2 T41704	hypothetical prote
13	181	10.4	540	2 T12704	leucine-rich prote
14	180.5	10.3	277	2 T30947	rsu-1 homolog - hu
15	180.5	10.3	572	2 T30947	hypothetical prote
16	180.5	10.3	613	2 A88684	hypothetical prote
17	179	10.3	559	2 T42998	protein AC7.2 [imp
18	177	10.1	268	2 T45616	Ras-binding protei
19	175	10.0	1839	1 OYBYK	hypothetical prote
20	173	9.9	424	2 S27783	adenylate cyclase
21	173	9.9	902	2 T00588	hypothetical prote
22	172.5	9.9	1119	2 AD1822	hypothetical prote
23	171.5	9.8	404	2 T08549	leucine-rich-repea
24	171	9.8	382	2 T04260	hypothetical prote
25	170	9.7	1039	2 T22117	hypothetical prote
26	166.5	9.5	2493	2 A55481	adenylate cyclase
27	166	9.5	268	2 T19697	hypothetical prote
28	165	9.5	1268	2 A49674	flightless-1 homol
29	163.5	9.4	2145	2 JC4747	adenylate cyclase

30	161.5	9.3	626	2 AE0123	probable antigenic
31	161.5	9.3	853	2 T17461	disease resistance
32	160.5	9.2	904	2 T46170	disease resistance
33	159.5	9.1	431	2 T27904	hypothetical prote
34	159	9.1	227	2 T27905	hypothetical prote
35	156.5	9.0	613	2 T15489	hypothetical prote
36	155	8.9	1029	2 T00712	protein kinase hom
37	154	8.8	1066	2 T15864	hypothetical prote
38	153.5	8.8	307	2 S36779	ribosome-binding p
39	153.5	8.8	371	2 T49908	hypothetical prote
40	153.5	8.8	1068	2 H96769	hypothetical prote
41	153.5	8.8	1232	2 T05322	hypothetical prote
42	153	8.8	1088	2 E86312	hypothetical prote
43	152	8.7	967	2 T48210	FlilA6.9 protein -
44	152	8.7	1011	2 C84524	probable disease r
45	151.5	8.7	720	2 T02361	hypothetical prote

ALIGNMENTS

RESULT 1

S45361
LR47 protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text_change 21-Jul-2000
C:Accession: S45361; S44046
R:Ntwasa, M.; Buchanan, S.G.S.C.; Gay, N.J.
Biochim. Biophys. Acta 1218, 181-186, 1994
A>Title: Sequence and expression of LR47, a novel embryonic leucine rich repeat protein
A:Reference number: S45361; MUID:94289476; PMID:8018718
A:Accession: S45361
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-428 <NTS>
A:CROSS-references: EMBL:X75760; NID:g415946; PIDM:CAA53387.1; PID:g415947
A:Gene: FlyBase:Lrr47
A:CROSS-references: FlyBase:FBgn0010398
C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology

Query Match.	16.2%;	Score	283.5;	DB 2;	Length	428;			
Best Local Similarity	27.7%;	Pred. No.	6.6e-14;						
Matches	92;	Conservative	55;	Mismatches	140;	Indels	45;	Gaps	8;
QY	15	LFQAISSSLKGFLSAMRLAHRG-----	CNVDTPVSTLTLPKTSFENFKCKWVITSKDD	69					
Db	31	LLIKCDPIQLKGFQTLKLGMDKDAINRLNINATAIPKQAP-----	QVRWVTSKRSSE	146					
QY	70	YPLSKNFPYSLEHLQTSYCGLVYDMRMMLCKLSRKLDSLHNHKKLPATIGDLHLQEL	129						
Db	147	YPI-KGFPRTLKSLTINNSQLVKLSFEICTLRNLTKLDVSGNKLAKIPSELGRU-PLTSL	234						
QY	130	NINNEHLSFS-----	VALCHSTLQKLSRLSLDSKKIKALPVQFCQLQELKMLKDD	182					
Db	205	HLGNLLGTQNDWCMLRGTKLC-----	QSLGELDLSGNGLTYFPPLPVKFFSLVSLNIN	259					
QY	183	NELIQPKCKIGQLINLRLFLSAARNKLPFLPSEFNLSLEYLDLQNTFE-----	QPKV	235					
Db	260	NLLSRPLFPATRMKALKLYVCSNELSESJPAVEDLRIDLDLVGNCSEKFNAAQKMC	319						
QY	236	LPVTKLQAPLTLLSSARTILHNRIPYSSHILPHLCQDLDTAKICVGRFLNS-----	290						
Db	320	LQKAASNSQPLWLLGARAVDKYMLPLSAGSIPAVLIDLIREAPRCPCGELCYAQKEDL	379						
QY	291	-----	FIQTTTNKLSVAHTVVVDN	312					
Db	380	FORVVPQKFSVKNLTYSREHQIYADVVLCD	411						

RESULT 2

T33295
hypothetical protein F26D11.i1 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T33295
R;Geisel, C.; Bradshaw, H.
submitted to the EMBL Data Library, May 1998
A;Description: The sequence of C. elegans cosmid F26D11.
A;Reference number: Z21316
A;Accession: T33295
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-717 <GEI>
A;Cross-references: EMBL:AF068716; PIDN: AAC37752.1; GSPDB: GN00023; CESP: F26D11.11
A;Experimental source: strain Bristol N2; clone F26D11
C;Genetics:
A;Gene: CESP:F26D11.11
A;Map position: 5
A;Introns: 34/2; 119/2; 226/1; 302/3; 368/3; 576/2; 623/2

Query Match 12.5%; Score 218; DB 2; Length 717;
Best Local Similarity 39.4%; Pred. No. 1.1e-08;
Matches 61; Conservative 17; Mismatches 69; Indels 8; Gaps 3;

QY 80 LEHLQSYCGLVVDVMDMLCKLSRLKLDLSNHHKIKLPATIGLIHLOELNADNHLESF 139
DB 38 LEDLNLTMMNKKLDHRLRILVDSDNELAVLPABIGNLTQIETLNARN----- 92
QY 140 SVALCHSTLQ--KSLRSLDSLSKNIKALPVQFCQLQELKLNKLDNDNELIQFPCKIGQLIN 197
DB 93 SIAPKPTQCNCKLLTTLNLSNPPFTLPETICECSITLSNETSLTLLPNSIGSLTN 152
QY 198 LRPLSARKKLPPLPSEFRLN-SLEVDLFGNTE 231
DB 153 LRVLARDNLLRTIPLSIVELRKLDELQONELE 187

RESULT 3
T31434
densin-180 - rat
N;Alternate names: postsynaptic density protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 07-Dec-1999
C;Accession: T31434
R;Apperson, M.L.; Moon, I.S.; Kennedy, M.B.
J. Neurosci. 16, 6839-6852, 1996
A;Title: Characterization of densin-180, a new brain-specific synaptic protein of the O-
A;Reference number: Z21033; MUID: 96421675; PMID: 8824323
A;Accession: T31434
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1495 <APP>
A;Cross-references: EMBL:U66707; NID: g1657757; PID: g1657758; PIDN: AAC52881.1
A;Experimental source: strain Sprague Dawley
C;Function:
A;Description: supposed to participate in specific adhesion between presynaptic and post

Query Match 11.9%; Score 208.5; DB 2; Length 1495;
Best Local Similarity 34.1%; Pred. No. 1.5e-07;
Matches 58; Conservative 28; Mismatches 75; Indels 9; Gaps 4;

QY 75 NFPYSLEHLQTSYCGLVVDVMDMLCKLSRLKLDLSNHHKIKLPATIGLIHLOELNNDN 134
DB 48 NEETLEELYLDANQIEELPKQFNQALRKLSIPNDLSLSTASLVNKLDELISKN 107
QY 135 HLESF--SVALCHSTLQSLRSLDSLSKNIKALPVQFCQLQELKLNKLDNDNELIQFPCKI 192
DB 108 GVQEPENIKCC-----KCLTITEASVNPISKLPGFTQLNLNLTQLYLNDADFLELPANF 162
QY 193 GOLINLRFLSARNKLPPLPSEFRLN-SLEVDLFGNTEP-EQPKVLPVK 240
DB 163 GRVLKRLIRLENLHKTLPKSMHKLQALRELQDNGNEFSLEPVLQIQ 212

RESULT 4

S60461
gene flightless-I protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 19-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 2i-Jul-2000
C;Accession: S60461; T08425
R;de Couet, H.G.; Fong, K.S.K.; Weeds, A.G.; McLaughlin, P.J.; Miklos, G.L.G.
Genetics 141, 1049-1059, 1995
A;Title: Molecular and mutational analysis of a gelsolin-family member encoded by the fly
A;Reference number: S60461; MUID: 96129280; PMID: 8582612
A;Accession: S60461
A;Molecule type: DNA
A;Residues: 1-1256 <DEC>
A;Cross-references: GB:AF017777; EMBL:U28044; NID: g3004652; PIDN: AAC28407.1; PID: g3004666
R;Maliszka, R.; de Couet, H.G.; Miklos, G.L.
Proc. Natl. Acad. Sci. U.S.A. 95, 3731-3736, 1998
A;Title: Data transferability from model organisms to human beings: insights from the fur
A;Reference number: Z16415; MUID: 98188272; PMID: 9520435
A;Accession: T08425
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1256 <MAL>
A;Cross-references: GB:AF017777; GB:U28044; NID: g3004652; PIDN: AAC28407.1; PID: PII
C;Genetics:
A;Gene: fli-I; flightless-I
A;Cross-references: FlyBase:FBgn0000709
A;Introns: 18/3; 1070/2; 1120/3
C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;53-75/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;76-100/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;101-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;124-147/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;148-170/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;171-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;220-242/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F;243-265/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F;266-288/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>
F;497-830/Domain: gelsolin repeat homology <GEL1>
F;892-1250/Domain: gelsolin repeat homology <GEL2>

Query Match 11.5%; Score 201.5; DB 2; Length 1256;
Best Local Similarity 33.7%; Pred. No. 3.9e-07;
Matches 63; Conservative 30; Mismatches 63; Indels 31; Gaps 6;

QY 75 NFPYSLEHLQTSYCGLVVDVMDMLCKLSRLKLDLSNHHKIKLPATIGLIHLOELNNDN 134
DB 210 NFPYSIDSLAN-----LC-----ELDSLNSLPKLPDCVYVNVTLVRLNSDN 252
QY 135 HLESFVALCHSTLQSLRSLDSLSKNIKALPVQFCQLQELKLNKLDNDNEL--IQFPCKI 192
DB 253 ELTELTAGV---ELWQRLSLSNLSRNQALVALPAALCKLPKLRLELVNDKLNFEIPSGI 309
QY 193 GOLINLRFLSARNKLPPLPSEF-RNLSLEYLDLFGNTEPQPKVLPVKIQALPTLLES 251
DB 310 GKLGALVFSANLLEWPEGCRCGALKQLNSCNR-----LITLPDAILHLESG 361
QY 252 ARTILEN 258
DB 362 DQDLN 368

RESULT 5
F86263
hypothetical protein F3K23.23 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C;Accession: F86263
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

QY 144 ---CHSTLQKSLRSLDLSKXKIKALPVQ-FCQQLKLNKLDNELLQFPCKIQQLNLR 199
 Db 1031 SISFKDFPKNMTSLTNKAQLSSIPGELLTKLSFLEKLELNQNNLTRLQPEISKLTJLV 1090
 QY 200 FLSSAANKLPFLSEPNL-SLEVLDFGNTF---EQKVLVPKIKOAPLTL-BSSAR 253
 Db 1091 FLSSAANKLPFLSEPNL-SLEVLDFGNTF---EQKVLVPKIKOAPLTL-BSSAR 253
 QY 254 TILNRIPIYSGSHI---IPFHLQCD 274
 Db 1151 NSFYHNSYGSLSKSLMFIAD 1174
 RESULT 8
 T01281
 N: probable leucine-rich-repeat protein At2g19330 [imported] - Arabidopsis thaliana
 C: Species: Arabidopsis thaliana (mouse-ear cress)
 C: Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 16-Feb-2001
 C: Accession: T01281; D84575
 R: Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
 euss, D.; Niemman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A: Title: Sequence and analysis of chromosome II BAC F27F23 genomic sequence.
 A: Reference number: Z14177
 A: Accession: T01281
 A: Status: translated from GB/EMBL/DBJ
 A: Molecule type: DNA
 A: Residues: 1-380 <ROU>
 A: Cross-references: EMBL:AC03058; NID:g3135250; PID:g3135263
 A: Experimental source: cultivar Columbia
 R: Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
 euss, D.; Niemman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A: Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A: Reference number: A84420; MUID:20083487; PID:10617197
 A: Accession: D84575
 A: Status: preliminary
 A: Molecule type: DNA
 A: Residues: 1-380 <STO>
 A: Cross-references: GB:AE002093; NID:g3135263; PIDN:AAC16463.1; GSPDB:GN00139
 C: Genetics:
 A: Gene: At2g19330; F27F23.13
 A: Map position: 2
 A: Introns: 96/3; 153/2
 Query Match 10.9%; Score 191; DB 2; Length 380;
 Best Local Similarity 32.5%; Pred. No. 5.4e-07;
 Matches 65; Conservative 28; Mismatches 75; Indels 32; Gaps 6;
 QY 65 TSKOYPLSKNPPYSLEHLQTSYCGILVRVDMRMCLKSLKLDLSHNNH-----112
 Db 47 SSPKSPSFNNBERLEVNLSGMALESPLNLAQICKLDLSNNHLQTIPESLTARL 106
 QY 113 -----IKKLIPATIGDLIHLQELNLDNHLSEPSVALCHSTLQKSLRSLDLSKN 160
 Db 107 LNLIALDVHNSQIKALPNSIGCLSKLTKLVNSGFLVSPFKSIQHC---KSLBELNANFN 163
 QY 161 KIKALP--VQFCOLQELKLNKLDNELLQFPCKIQQLNLRFLSAAARNKLPFLSEPNL 218
 Db 164 KLRLPDSIGF-ELTNRLKLSINSKLSLPSITHTLSLRVLDARLNCMLIPDDLENL 222
 QY 219 -SLEYLDLFGNTFEQKVLVP 237
 Db 223 INLEILNVSON-FQYLSALP 241
 RESULT 9
 C84552
 N: hypothetical protein At2g17440 [imported] - Arabidopsis thaliana
 C: Species: Arabidopsis thaliana (mouse-ear cress)
 C: Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C: Accession: C84552

R: Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
 euss, D.; Niemman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A: Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A: Reference number: A84420; MUID:20083487; PID:10617197
 A: Accession: C84552
 A: Status: preliminary
 A: Molecule type: DNA
 A: Residues: 1-526 <STO>
 A: Cross-references: GB:AE002093; NID:g2494144; PIDN:AAB86525.1; GSPDB:GN00139
 C: Genetics:
 A: Gene: At2g17440
 A: Map position: 2
 Query Match 10.8%; Score 189.5; DB 2; Length 526;
 Best Local Similarity 30.6%; Pred. No. 1.1e-06;
 Matches 77; Conservative 27; Mismatches 97; Indels 51; Gaps 10;
 QY 77 PYSLEHLQTSYCGILVRVDMRMCLC-----LKSRLKLDLSHNNKIKLPATIGDLIHL 126
 Db 223 PDSLGKLS-----LVRDLSENCIMVLPATIGGLISLTRLDLHSNRIGQLPESIGDLNL 278
 QY 127 QELNLDNHLSEPSVA---LCHSTLQKSLRSLDLSKXKIKALPVQFCOLQELKLNKLDN 183
 Db 279 VNLNLSGNQLSSLPSSFNRLIH-----LEELDSSNSLSILPESIGSLVSLKLDVETN 332
 QY 184 ELIQFPCKIQQLNLRFLSAAARNKLPFLSEPNL-SLEVLDFGNTFEQKVLVP-VIKL 241
 Db 333 NIEIPIHSISGCSWEELRADYNRLKALPEAVGKLSLEILTTRYNNIRQ---LPTWSS 389
 QY 242 QAPLTLLLESSARTILENRIPIYSGSHIIPFHLQCDLTAIKVCGRFLCNLSFIQGTITMNLH 301
 Db 390 MANLKELDVSEFNL-----ESVPESICYAKTLVXLNI-----GNNFANLR 429
 QY 302 SVAHTVVLVDNL 313
 Db 430 SLFG---LIGNL 438
 RESULT 10
 A88536
 N: protein B0523.5 [imported] - Caenorhabditis elegans
 C: Species: Caenorhabditis elegans
 C: Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C: Accession: A88536
 R: anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A: Title: Genome sequence of the nematode C. elegans: a platform for investigating biology,
 A: Reference number: A75000; MUID:99069613; PMID:9851916
 A: Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.eleg
 A: Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A: Accession: A88536
 A: Status: preliminary
 A: Molecule type: DNA
 A: Residues: 1-1257 <STO>
 A: Cross-references: GB:chr_III; PIDN:AAB37088.1; PID:g1695248; GSPDB:GN00021; CESP:B0523.
 C: Genetics:
 A: Gene: B0523.5
 A: Map position: 3
 Query Match 10.8%; Score 189; DB 2; Length 1257;
 Best Local Similarity 25.3%; Pred. No. 3.5e-06;
 Matches 82; Conservative 46; Mismatches 100; Indels 100; Gaps 12;
 QY 14 LRFQALSSSLKGLSAMRLAHRGCVNTPVSTLTLPKTSFEPFNPKMTKMTSKKDYPLS 73
 Db 6 LQFVKGIDFSGNDF-SGDRFPH-----DVEQMTQMTWLKNDLSLEQV-----PDE 50
 QY 74 KNFPYSLEHLQTSYCGILVRVDMRMCLKSLRK-----LDL 108
 Db 51 LSRCANLEHLQAHNQILISVEGELSRLPRLSRVIVRDNLNKTAGITPTDPRMKDLTIIDL 110

A;Introns: 199/1; 503/3
A;Note: F15J1.40

```

Query Match      10.4%; Score 181.5; DB 2; Length 549;
Best Local Similarity 32.8%; Pred.No 4.5e-06;
Matches 62; Conservative 24; Mismatches 86; Indels 17; Gaps 6;

QY 100 LKSLRKLDSLHNHIKKLPATIGDLIHLQELNLDNHNLEGSFVALCHSTLQKSLRSIDLK 159
Db 268 LSSITKLLDLSNRIGQLPESIGELLNVILNLSGNQLSLPSAF--SRIVR-LEELDLSL 324

QY 160 NKIKALPVQCOLQELKNLKLDDNELIQPCKTGQLINLRFLLSAARNKLPELSEFRNL- 218
Db 325 NNDPIIPESIGSVLSKKLDVFNDDLEEPYISGGCSSLTELADYNKRALPEAIGKIT 384

QY 213 SLEYLDLFGNTFQPKVLPVLIKQAPLTLLESSARTILHNRIPYGS-HIPIFFHLCODLDT 277
Db 385 TLEILSVRYNNIR-----QLPTTM--SSLASIKELDVSENELESVPESLCFATTL 432

QY 278 AKICVCGRF 286
Db 433 VKLNIGNNF 441

RESULT 13
T12704
leucine-rich protein - common sunflower
C;Species: Helianthus annuus (common sunflower)
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 13-Aug-1999
C;Accession: T12704
R;Evvard, J.L.
submitted to the EMBL Data Library, September 1994
A;Reference number: Z17575
A;Accession: T12704
A;Status: preliminary; translated from GS/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-540 <EVR>
A;Cross-references: EMBL:X81997; NID:e977240; PID:e120583
C;Genetics:
A;Gene: SPL7

Query Match      10.4%; Score 181; DB 2; Length 540;
Best Local Similarity 29.0%; Pred.No. 4.8e-06;
Matches 90; Conservative 42; Mismatches 122; Indels 56; Gaps 15;

QY 12 LQLLFTQALSSSLKGFSLSAHLRGNCVDTPVSTLTP-----VKTSEFENFKTKMVIIS 66
Db 166 LQKLDYTKSA---GKLSAL-----DKRVSSMPTIAEEVK-EEMADILQDAMNG 212

QY 67 KXDYPLS-KNFFY-----SLEHLOTSYGLRVDMRMMLCLSKRLKLDLSHNHIKKL 116
Db 213 VERIDLSRRRLPFVPEAFGKLHKTLSLDSSNKLTAIPESLAGLTSLEELNLSANLFESL 272

QY 117 PATIGDLIHLQELNLDNHNLEGSFVALCHSTLQKSLRSLDKNKIKALPVQF-CQLQEL 175
Db 273 PDTGISLQHLQELNVRNKLTLSDPGICKC---RSLLEDASFNOITYLPANIGYGLNL 329

QY 176 XNKLDDNELIQPKIGKIGLINLRFLSAARNKLPELSEFRNL-SLEYLDLFGNTFQPK 234
Db 330 KKLINPLNNVRSLPTSIGEMISLOVLDVHNTLGRGIPPSIGMKLKEVLNL-GSNFNDFT 388

QY 235 VLP-----VKKL-----QAPLTL--LESSARTIL-HNRIPYGSHIIPFHLCOQL 275
Db 389 ALPETIGSTLELRLELDICNNQIQLPITFGRVLSLRLVWDENPLTVSP---PEVVAEGV 445

QY 276 DTAKICVCGR 285
Db 446 EAVKVMYMKR 455

RESULT 14
I60122
isp-1 homolog - human
```

Search completed: June 15, 2004, 19:03:20
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: June 15, 2004, 10:01:25 ; Search time 48 Seconds
(without alignments)
1983.821 Million cell updates/sec

Title: US-10-009-557-9

Perfect score: 1745

Sequence: 1 MYV-SPVEFIILQLFIQAI.....PIISYFCSLGCYVNSDMLK 338

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1158786 seqs, 281726120 residues

Total number of hits satisfying chosen parameters: 1158786

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/ECT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/ECTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1649	94.5	416	15	US-10-312-088-44
2	1315	75.4	259	15	US-10-312-088-43
3	227.5	13.0	444	15	US-10-369-493-20189
4	218	12.5	717	15	US-10-369-493-6391
5	217	12.4	1489	15	US-10-115-482-18
6	214	12.3	231	15	US-10-369-493-11177
7	201.5	11.5	258	15	US-10-369-493-19830
8	199.5	11.4	365	12	US-10-087-192-69
9	197.5	11.3	546	9	US-09-903-320-250
10	197.5	11.3	546	9	US-09-903-068B-250
11	197.5	11.3	546	9	US-09-905-291A-250
12	197.5	11.3	546	9	US-09-902-853-250
13	197.5	11.3	546	9	US-09-907-824-250
14	197.5	11.3	546	9	US-09-907-841-250
15	197.5	11.3	546	10	US-09-904-011-250

16	197.5	11.3	546	10	US-09-906-742-250	Sequence 250, App
17	197.5	11.3	546	10	US-09-906-838-250	Sequence 250, App
18	197.5	11.3	546	10	US-09-907-613-250	Sequence 250, App
19	197.5	11.3	546	10	US-09-907-942-250	Sequence 250, App
20	197.5	11.3	546	10	US-09-904-859-250	Sequence 250, App
21	197.5	11.3	546	10	US-09-909-204-250	Sequence 250, App
22	197.5	11.3	546	10	US-09-904-820-250	Sequence 250, App
23	197.5	11.3	546	10	US-09-904-786-250	Sequence 250, App
24	197.5	11.3	546	10	US-09-906-646-250	Sequence 250, App
25	197.5	11.3	546	10	US-09-906-700-250	Sequence 250, App
26	197.5	11.3	546	10	US-09-903-786-250	Sequence 250, App
27	197.5	11.3	546	10	US-09-902-903-250	Sequence 250, App
28	197.5	11.3	546	10	US-09-903-749A-250	Sequence 250, App
29	197.5	11.3	546	10	US-09-904-119-250	Sequence 250, App
30	197.5	11.3	546	10	US-09-904-956-250	Sequence 250, App
31	197.5	11.3	546	10	US-09-902-736-250	Sequence 250, App
32	197.5	11.3	546	10	US-09-907-794-250	Sequence 250, App
33	197.5	11.3	546	10	US-09-903-943-250	Sequence 250, App
34	197.5	11.3	546	10	US-09-904-462-250	Sequence 250, App
35	197.5	11.3	546	10	US-09-907-925-250	Sequence 250, App
36	197.5	11.3	546	10	US-09-902-692-250	Sequence 250, App
37	197.5	11.3	546	10	US-09-903-520-250	Sequence 250, App
38	197.5	11.3	546	10	US-09-905-056-250	Sequence 250, App
39	197.5	11.3	546	10	US-09-909-064-250	Sequence 250, App
40	197.5	11.3	546	10	US-09-904-553-250	Sequence 250, App
41	197.5	11.3	546	10	US-09-905-381-250	Sequence 250, App
42	197.5	11.3	546	10	US-09-905-088-250	Sequence 250, App
43	197.5	11.3	546	10	US-09-907-575-250	Sequence 250, App
44	197.5	11.3	546	10	US-09-905-075-250	Sequence 250, App
45	197.5	11.3	546	10	US-09-902-759-250	Sequence 250, App

ALIGNMENTS

RESULT 1

US-10-312-088-44

; Sequence 44, Application US/10312088

; Publication No. US20030219862A1

; GENERAL INFORMATION:

; APPLICANT: Agarwal, Pankaj

; APPLICANT: Cogswell, John P.

; APPLICANT: Kabric, Karen S.

; APPLICANT: Lai, Ying-Ta

; APPLICANT: Martensen, Shelby A.

; APPLICANT: Murdock, Paul R.

; APPLICANT: Smith, Randall F.

; APPLICANT: Strum, Jay C.

; APPLICANT: Xiang, Zhaoying

; APPLICANT: Xie, Qing

; APPLICANT: Rizni, Safia K.

; TITLE OF INVENTION: NOVEL COMPOUNDS

; FILE REFERENCE: GP50029

; CURRENT APPLICATION NUMBER: US/10/312,088

; CURRENT FILING DATE: 2002-12-20

; PRIOR APPLICATION NUMBER: PCT/US01/19929

; PRIOR FILING DATE: 2001-06-22

; PRIOR APPLICATION NUMBER: 60/213,161

; PRIOR FILING DATE: 2000-06-22

; PRIOR APPLICATION NUMBER: 60/213,156

; PRIOR FILING DATE: 2000-06-22

; NUMBER OF SEQ ID NOS: 44

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 44

; LENGTH: 416

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-312-088-44

Query Match 94.5%; Score 1649; DB 15; Length 416;
Best Local Similarity 99.1%; Pred. No. 8,5e-154;
Matches 320; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/296 965

```
Query Match      12.3%; Score 214; DB 15; Length 231;  
Best Local Similarity   37.5%; Pred. No. 1.3e-12;  
Matches        66; Conservative    33; Mismatches    67; Indels     10; Gaps       5;
```

	QY	DB	LKSLRKLDLSHNHKKLPATIGDLIHLQLCLNNDNHLESFVALCHSTTLQKSRLSRLDLSK	159
	:	:	: :: :::	:: :::
48	LKNLKTLDLSGNQLTQPSEIGLNKLNTLINLYDNQTOLTQPEIKE---	LKNLTALTFN	104	
	:	:	: :: :::	:: :::
160	NKIKAIPVQFCOLQEMLNKLDNELIQPCIKGVLINLRFTSAARNKLPFLPFEFRNL-	218		
	:	:	: :: :::	:: :::
105	NXLTIQPIEIGLKLNLETIIYCNGLTQTQPBEIGELKNLSIALNNKULTQTPSEIGNLK	164		
	:	:	: :: :::	:: :::
219	SLEYILDLCGTTEQKVLP--VIKQAPUTLLESSARTILNHRIDPYGSHTPFHL	271		
	:	:	: :: :::	:: :::

DB 165 NLETSLYRNQLE---LPPEIGKLENKATLYIDNNKLTLPBPSELSKMLITLNL 217

RESULT 7

US-10-369-493-19830
; Sequence 19830, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19830
; LENGTH: 288
; TYPE: PRT
; ORGANISM: No. US20030233675altoc punctiforme
US-10-369-493-19830

Query Match 11.5%; Score 201.5; DB 15; Length 288;
Best Local Similarity 33.0%; Pred. No. 2.9e-11;
Matches 61; Conservative 29; Mismatches 78; Indels 17; Gaps 6;
QY 64 ITSXKDYPLSKN-----FP---YSLEHLQ---SYGGLVVRVDMRLCLKSLRKLDLSHNHI 113
DB 32 LTQQLDLSRNRQVQLPEAASLTQLQTLDSNNKLTQLPEAASLTARLQRLDLSNNOL 91
QY 114 KQIPATIGDLHQLNLDNHNLESFVALCHSTLQKSLRSLDLSKXKIKALPVQFCQLQ 173
DB 92 TELPEAASLTALQQLNLRNNQLTELPEAASLT---RLQRLDLSNNQUTELPEAASLT 148
QY 174 ELKNLKLDNELIQFCKIGQILNRLFLSAARNKLPFLPSEPNL-SLEYLDLFGNTPEQ 232
DB 149 QQQSFDLSHNLTELPSLSRLLYLEIFDCGNSNLRQVPSVIKELGKELYIYANDLE- 207
QY 233 PKVLP 237
DB 208 --VIP 210

RESULT 8

US-10-087-192-69
; Sequence 69, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(365)

; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-087-192-69

Query Match 11.4%; Score 199.5; DB 12; Length 365;
Best Local Similarity 28.6%; Pred. No. 6.5e-11;
Matches 77; Conservative 29; Mismatches 88; Indels 75; Gaps 8;
QY 72 LSKNPFYSLEHLQTSYCGLVVRVDMRLCLKSLRKLDLSHNHIKKLPATIGDLHLQLNL 131
DB 55 LSKN-----RLVEVPMELCQFVSEILNLNHNIRVIRPEAVNLQMLCHLNL 101
QY 132 NDNHLESFVALCHSTLQKSLRSLDLSKXKIKALPVQFCQLKLDLDDNELIQFCK 191
DB 102 SRNLQSLPACLC---GHLKVLTIASNNKLSLPEEIGQLKQLMELDVSCNEITALPQ 157
QY 192 IGOINLRFLSAARNKLPFLPSEPNLSLEYLDLFGNTPEQKVLPEVI---KLQAPUTLL 248
DB 158 IGOILSLRELNVRRNLYKVLPELVDLPLVKDFSCN---KVLVIVPCFREMQLQVLL 214
QY 249 ESSARTILHNRIPIYGHSHIIPHLQDLOLTAIKVCGRFCFLNSFIQTTTNNLHSHVAHV 308
DB 215 ENN-----PLQS-----PPAQICTKGK-----VH----- 233
QY 309 IVDNLGCTEAPIISYFCSLGCYVNSDML 337
DB 234 -----IFKYLSTQACQIKTSDSL 251

RESULT 9

US-09-909-320-250
; Sequence 250, Application US/09909320
; Patent No. US20020132240A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Deshoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909,320
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/43,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/445,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/446,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08

PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 250
LENGTH: 546
TYPE: PRT
ORGANISM: Homo Sapien
US-09-909-320-250

Query Match 11.3%; Score 197.5; DB 9; Length 546;
Best Local Similarity 29.1%; Pred. No. 1.8e-10;
Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;
QY 57 NFKTKWITSKKYPLSKNFFYSLEHQTSGYGLVRVDMRLCLSKSLKLDLS----- 109
Db 312 NEGTKLVIN-----SLKXMANLTELRLRCDLERIPHSFSLNQLQIDLKNNLXTI 365
QY 110 -----HNHIKLPATIGDILHIQELNLDNHNLESFSAVLCSTLOKS 151
Db 366 EELISFQHLRLTCLKWNHIAVPIQIGNLTNRLYLNENKIEKPTQFYC---RK 422
QY 152 LRSDLSSNKIKALPVQFCQQLKLNLDONELIQFCKIGQLINLRLSARKNLPL 211
Db 423 LRYLDLSHNNLTFLPADIGLLQNLQNLAITANRIETLPPELFCCKRLALHGLNVLQSL 482
QY 212 PSEFRLN-SLEYDLFGNTFFQPKVLPVYKQLAPLTLESSARTI---LHNRIP 261
Db 483 PSRVGELTNLTQIELRGNLE---CLPVELGECF--LIXRSLGVVEEDLFTNLP 531

RESULT 10
US-09-909-088B-250
Sequence 250. Application US/0909088B
Patent No. US20020146709A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary B.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.

APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/909/088B
CURRENT FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 250
LENGTH: 546
TYPE: PRT
ORGANISM: Homo Sapien
US-09-909-088B-250

Query Match 11.3%; Score 197.5; DB 9; Length 546;
Best Local Similarity 29.1%; Pred. No. 1.8e-10;
Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;
QY 57 NFKTKWITSKKYPLSKNFFYSLEHQTSGYGLVRVDMRLCLSKSLKLDLS----- 109
Db 312 NEGTKLVIN-----SLKXMANLTELRLRCDLERIPHSFSLNQLQIDLKNNLXTI 365
QY 110 -----HNHIKLPATIGDILHIQELNLDNHNLESFSAVLCSTLOKS 151
Db 366 EELISFQHLRLTCLKWNHIAVPIQIGNLTNRLYLNENKIEKPTQFYC---RK 422
QY 152 LRSDLSSNKIKALPVQFCQQLKLNLDONELIQFCKIGQLINLRLSARKNLPL 211
Db 423 LRYLDLSHNNLTFLPADIGLLQNLQNLAITANRIETLPPELFCCKRLALHGLNVLQSL 482
QY 212 PSEFRLN-SLEYDLFGNTFFQPKVLPVYKQLAPLTLESSARTI---LHNRIP 261

DB 483 PSRVGELTNLTQIELRGNRL---CLPVELGECF--LLKRSGLVBEEDLFTLTP 531

RESULT 11

US-09-905-291A-250

; Sequence 250, Application US/09905291A

; Patent No. US20020160374A1

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnovers, Luc

; APPLICANT: Ertan, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth, J.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Mather, Jennie P.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tomas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/905,291A

; CURRENT FILING DATE: 2001-07-12

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: PCT/US99/28214

; PRIOR FILING DATE: 1999-11-29

; PRIOR APPLICATION NUMBER: PCT/US99/28313

; PRIOR FILING DATE: 1999-11-30

; PRIOR APPLICATION NUMBER: PCT/US99/28564

; PRIOR FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: PCT/US99/28565

; PRIOR FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: PCT/US99/30095

; PRIOR FILING DATE: 1999-12-16

; PRIOR APPLICATION NUMBER: PCT/US99/30911

; PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US99/30999

; PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US00/00219

; PRIOR FILING DATE: 2000-01-05

; NUMBER OF SEQ ID NOS: 423

; SEQ ID NO 250

; LENGTH: 546

; TYPE: PRT

; ORGANISM: Homo Sapien

US-09-905-291A-250

Query Match 11.3%; Score 197.5; DB 9; Length 546;

Best Local Similarity 29.1%; Pred. No. 1.8e-10;

Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;

QY 57 NFKTRVITSKDYPPLSKNPPYSLEHLQTSYCGVVRVDMELCLKSKRLDLS----- 109

DB 312 NEGTKLIVLN-----SLKFCANLTELRLCRLERIPHSIFSLHNLQETDLKNNLKITI 365

QY 110 -----HNHKKLPATIGDLIHLQELNLDNHLSEFSVALCHSTLQKS 151

DB 366 EEIISFQHLHRLTCLKLWYXHYAVIPIQIGNLNLRLYLNRRKIKIPTQLFVC---RK 422

QY 152 LRSLDLSKNIKALPVQFCQLQELKMLKDDNELIOFPCKIGQLINIRFLSAARKLPFL 211

DB 423 LRYLDLSHNNLTPLPADIGLLQNLQNLATITANRIETLPPFLFOCKRLALHGLGNVLQSL 482

QY 212 PSFRNL-SLEYLDLRCNTFEQPKVLPTKQAPLTCLESSARTI---LHNRIP 261

DB 483 PSRVGELTNLTQIELRGNRL---CLPVELGECF--LLKRSGLVBEEDLFTLTP 531

RESULT 12

US-09-902-853-250

; Sequence 250, Application US/09902853

; Publication No. US20020192659A1

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnovers, Luc

; APPLICANT: Ertan, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth, J.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Mather, Jennie P.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tomas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/902,853

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: US/09/665,350

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

```

; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 250
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-902-853-250

Query Match      11.3%; Score 197.5; DB 9; Length 546;
Best Local Similarity 29.1%; Pred. No. 1.8e-10;
Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;

Qy 57 NFKTKMVTSSKKDYPLSKNFPYSLEHLQTSYGLVVRVDMRMCLCKSLRKLDLS----- 109
Db 312 NEGTKLIIVN-----SLKMANLTELRLCRLSRIPIHSHFSLHNLQELDKNNLKTI 365
Qy 110 -----HNHKKLPATIGDLIHQELNLDNDNHLESFVALCHSTLOKS 151
Db 366 EEIISFQHLHRLTCLKWNHIAIPIQIGNLTNLERLYLNRNKIEKPTQLFYC---RK 422
Qy 152 LRSLDLKNKIKALPVQFCQQLQELKNLXLDNELLQFCKTGQLINLRFLSAARNKLPFL 211
Db 423 LRYLDLSHNNLTFCPADGLLQNLCLNLAICANRIETLPELFCQCKLRAHLGNNVLQSL 482
Qy 212 PSEFNL-SLEYLDLFGNTFQPKVLPVVKLOAPLTLESSARTI---LHNRIP 261
Db 483 PSRVGELTNLTQIELRGNLE---CLPVELGEC2--LLKRSGLVVEEDLFNTLP 531
```

```

RESULT 13
US-09-907-824-250
; Sequence 250, Application US/09907824
; Publication No. US2002019767A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Deanoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
```

```

; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,824
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 250
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-907-824-250

Query Match      11.3%; Score 197.5; DB 9; Length 546;
Best Local Similarity 29.1%; Pred. No. 1.8e-10;
Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;

Qy 57 NFKTKMVTSSKKDYPLSKNFPYSLEHLQTSYGLVVRVDMRMCLCKSLRKLDLS----- 109
Db 312 NEGTKLIIVN-----SLKMANLTELRLCRLSRIPIHSHFSLHNLQELDKNNLKTI 365
Qy 110 -----HNHKKLPATIGDLIHQELNLDNDNHLESFVALCHSTLOKS 151
Db 366 EEIISFQHLHRLTCLKWNHIAIPIQIGNLTNLERLYLNRNKIEKPTQLFYC---RK 422
Qy 152 LRSLDLKNKIKALPVQFCQQLQELKNLXLDNELLQFCKTGQLINLRFLSAARNKLPFL 211
Db 423 LRYLDLSHNNLTFCPADGLLQNLCLNLAICANRIETLPELFCQCKLRAHLGNNVLQSL 482
Qy 212 PSEFNL-SLEYLDLFGNTFQPKVLPVVKLOAPLTLESSARTI---LHNRIP 261
```

Db 483 PSRVGELTNLTQIELRGNRL---CLPVELGECF--LLKRSGLVVEEDLNTLP 531

RESULT 14

US-09-907-841-250
; Sequence 250, Application US/09907841
; Publication No. US20020198366A1

GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Baton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Christopher J.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/907,841

; CURRENT FILING DATE: 2001-11-20

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: PCT/US99/28214

; PRIOR FILING DATE: 1999-11-29

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 423

; SEQ ID NO 250

; LENGTH: 546

; TYPE: PRT

; ORGANISM: Homo Sapien

US-09-907-841-250

Query Match

Best Local Similarity 11.3%; Score 197.5; DB 9; Length 546;

Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;

QY 57 NFKTQWITSKXDPYLSKKNFPYSLEHQTSCYGLVVDMMCLCKSLRKLDLS-----109

Db 312 NEGTKLIWN-----SLKKNANITELIRCDLERIPHSIFSLHNLQELDKNNLXTI 365
QY 110 -----HHHIKKLPATIGLIHLQELNLDNHSFVALCHSTLQKS 151
Db 366 ERIISFOHLRLTCLKWYHNIAYIPQIGNLTNLERLYNNRKIEKPTQLFYC---RK 422
QY 152 LRSIDLKNNIKALPVQFCQIQELKNLKDNDNELIQFPCKIGQILNRFLSAARNKLPFL 211
Db 423 LRYLDLSHNNITFLPADIGLLQNLQNLAITANRIETLPPELFQCKURALHGNVLOSL 482
QY 212 PSEFRNL-SLEYLDLFGNTFEQPKVLPVILKQAPLTLESARTI---LNNRIP 261
Db 483 PSRVGELTNLTQIELRGNRL---CLPVELGECF--LLKRSGLVVEEDLNTLP 531

RESULT 15

US-09-904-011-250

; Sequence 250, Application US/09904011

; Publication No. US20030003530A1

GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Baton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/904,011

; CURRENT FILING DATE: 2001-07-11

; PRIOR APPLICATION NUMBER: 09/665,350

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: PCT/US99/28214

; PRIOR FILING DATE: 1999-11-29

; PRIOR APPLICATION NUMBER: PCT/US99/28313

```

; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 250
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-904-011-250

Query Match      11.3%; Score 197.5; DB 10; Length 546;
Best Local Similarity 29.1%; Pred. NO. 1.8e-10;
Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;

QY 57 NFKTMMVTSKXDYPLSKNPPYSLEHLQTSYCGLVNRMMLCKLSRKLDLS----- 109
Db 312 NEGTXLIVLN-----SLKQMANLTELIRCDLERIPHSIFSLHNLQELDKDNNKTI 365
QY 110 -----HNHIKKLPATIGDLHLOELNNDNHLESFVALCHSTLQXS 151
Db 366 EBIIISFOHLRLTCLKLWYNHAYIPQIGNLTNLERLYNKNKIEKIPTQFYC---RK 422
QY 152 LRSLSKXNKIKALPVQFQQLQELKXNLKDDNELIQPCKIGQLINLRFISAARNKLPFL 211
Db 423 LRYLDLSHNNLTFLPADIGLLQNLQNLATANRIETLPPELPQCRKLRAHLGNVQLSL 482
QY 212 PSEFENL-SLEYLDLFGNTFQPKVLPVTKLQAPLTLLLESSARTI---LHNRIP 261
Db 483 PSRVGELTNLTQIELRGNRL---CLPVELGECF--LLKRSGLIVVEDLFTNIP 531

```

Search completed: June 15, 2004, 10:04:56
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2004, 10:00:10 ; Search time 23 Seconds
(without alignments)
758.677 Million cell updates/sec

Title: US-10-009-557-9

Perfect score: 1745

Sequence: 1 MVLSPVEFIIQLLFIQAI.....PIISYFCSLGLVYNSDMLK 338

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/prodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/2/iaa/6C COMB.pep.*
6: /cgn2_6/prodata/2/iaa/6D COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	197.5	11.3	546	4	US-09-907-794A-250
2	197.5	11.3	546	4	US-09-905-125A-250
3	197.5	11.3	546	4	US-09-902-775A-250
4	187.5	10.7	501	4	US-09-907-794A-185
5	187.5	10.7	501	4	US-09-905-125A-185
6	187.5	10.7	501	4	US-09-902-775A-185
7	179	10.3	559	4	US-09-081-149-2
8	178.5	10.2	582	4	US-09-081-149-8
9	177.5	10.2	582	4	US-09-081-149-7
10	162	9.3	1874	4	US-09-331-403-2
11	154	8.8	723	4	US-09-434-408-2
12	152	8.7	292	4	US-09-312-283C-384
13	148.5	8.5	560	3	US-08-592-500-2
14	148.5	8.5	560	3	US-08-195-006-2
15	148.5	8.5	560	3	US-09-063-950-4
16	148.5	8.5	560	5	PCT-US94-07644A-2
17	148.5	8.5	1240	3	US-08-930-996A-4
18	147	8.4	799	3	US-09-180-439-6
19	147	8.4	1112	3	US-09-353-585-2
20	147	8.4	1112	3	US-09-353-585-3
21	142.5	8.2	947	4	US-09-228-986-73
22	141.5	8.1	863	2	US-08-666-271-2
23	140.5	8.1	673	3	US-09-063-950-2
24	140.5	8.1	910	4	US-09-228-986-72
25	138.5	7.9	806	3	US-08-945-983-2
26	138.5	7.9	907	4	US-09-170-496D-264
27	138.5	7.9	907	4	US-09-170-496D-278

ALIGNMENTS

RESULT 1

US-09-907-794A-250
; Sequence 250, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Askenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547

28	138	7.9	999	2	US-08-473-553A-5
29	137	7.9	115	4	US-09-621-976-4266
30	136.5	7.8	1016	3	US-09-180-439-8
31	134.5	7.7	968	3	US-09-180-439-3
32	134.5	7.7	968	3	US-09-180-439-4
33	133	7.6	1012	2	US-08-475-891A-4
34	133	7.6	1025	2	US-08-567-375-4
35	133	7.6	1025	2	US-08-587-680A-4
36	132	7.6	1119	4	US-09-907-794A-294
37	132	7.6	1119	4	US-09-905-125A-294
38	132	7.6	1119	4	US-09-902-775A-294
39	131	7.5	320	1	US-07-613-083B-1
40	130.5	7.5	1525	3	US-09-191-647-2
41	130.5	7.5	1525	3	US-09-540-245A-2
42	130.5	7.5	1525	3	US-09-540-153-2
43	129.5	7.4	1059	4	US-09-907-794A-290
44	129.5	7.4	1059	4	US-09-905-125A-290
45	129.5	7.4	1059	4	US-09-902-775A-290

Sequence 5, Appli
Sequence 4266, Ap
Sequence 3, Appli
Sequence 3, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 24, App
Sequence 234, App
Sequence 1, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 230, App
Sequence 290, App
Sequence 290, App

```

RESULT 2
US-09-905-125A-250
  / Sequence 250, Application US/09905125A
  / Patent No. 6664376
  / GENERAL INFORMATION:
    / APPLICANT: Genentech, Inc.
    / APPLICANT: Ashkenazi, Avi.
    / APPLICANT: Botstein, David
    / APPLICANT: Desnoyers, Luc
    / APPLICANT: Eaton, Dan L.
    / APPLICANT: Ferrara, Napoleone
    / APPLICANT: Filvaroff, Ellen
    / APPLICANT: Fong, Sherman
    / APPLICANT: Gao, Wei-Qiang
    / APPLICANT: Gerber, Hanspeter
    / APPLICANT: Gerritsen, Mary E.
    / APPLICANT: Goddard, A.
    / APPLICANT: Godowski, Paul J.
    / APPLICANT: Grimaldi, Christopher J.
    / APPLICANT: Gurney, Austin L.
    / APPLICANT: Hillan, Kenneth, J.
    / APPLICANT: Kljavin, Ivar J.
    / APPLICANT: Mather, Jennie P.
    / APPLICANT: Pan, James
    / APPLICANT: Paoni, Nicholas F.
    / APPLICANT: Roy, Margaret Ann
  /

```

Query Match	11.3%	Score 197.5	DB 4	Length 546
Best Local Similarity	29.1%	Pred. NO. 3.5e-12		
Matches	68	Conservative 33	Mismatches 90	Indels 43
Gaps	7			

QY	57	NFTKMWITSKKDYPKSKNFFPSLEHLSQSYCGLVQVDMRMCLCKSLRKLDLS-----	109
DB	312	NEGTKLIIVN-----SLKKMANTELRLRCOLERPHSIFSLHNLQELDKDNNLKT	365
QY	110	-----HHHIXKLPTATIGDLHLQLQELNLDNHLSFVSVALCHSTLQKS	151
DB	366	EETISFQHLRLTCLKLYWNHAIYIPQIGNTLTNLERLYLRNKKIEKIPQLFYC--RK	422
QY	152	LRSLDLSKNNKIKALPVQFCQLQELKXLDNDNELIQPCKIGQLINLRLSAARNKL	211
DB	423	LRYLDSHNNLFLPADIGULQNLQALATANNRIETLPPELFCCKRUAHLGNVQLQS	482
QY	212	PSBFRL-SLEYLDLFGNTFQPKVLPIVKLQAPLTLLBSSARTI---LHNRIP	261
DB	483	PSRAGELTNLTQELRGNLE---CLPVELGECF--LLKRSQSWBEDLFTNLP	531

RESULT 3
US-09-902-775A-250

```

Query Match          11.3%; Score 197.5; DB 4; Length 546;
Best Local Similarity 29.1%; Pred. No. 3.5e-12;
Matches 68; Conservative 33; Mismatches 90; Indels 43; Gaps 7;

QY  57  NFKTKAVITSKKDYPLSKNFPYPKSYLEHLQTSYGLVRVDMRLCLKSLRKLDLS----- 109
Db    312  NECTKLIVLN-----SLAKMANLTELILRCDLERIPHSIFSLHNQLQIDLDKNNLKTI 365

QY  110  -----HHHKKLPATIGDLHLQELNLDNHLESFSVALCHSTLQKS 151
Db    366  EEIISFQHLRHRLTCLKLVNHYAVIDPIQNLNTERLYNRNKKIEKIPQLFYC---RK 422

QY  152  LRSLDLSKKIKALPQVQQLQELNKLDDNELIQFCKIGQLINLRFSLAARNKLPLFL 211
Db    423  LRYLDLSHNNLTFELPADIGLQNLQNLATANRIETLPPELFOCKLRALHLCNNVLQSL 482

QY  212  PSEFRNL-SLEVDLFGNTFEQPKVLPIKQAPLTLLESSARTI---LHNRIIP 261
Db    483  PSRVGLTNLQLELGNRL-----CLPVELGCEP--LLKRSGLVVEDLFTNLTIP 531

RESULT 4
US-09-907-794A-185
: Sequence 185, Application US/09907794A
: Patent No. 6635468
: GENERAL INFORMATION:
: APPLICANT: Genentech, Inc.
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan L.
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, A.
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, Christopher J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth, J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Mather, Jennie P.
: APPLICANT: Pan, James
: APPLICANT: Paoletti, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William, I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: ACIDS Encoding the Same
: FILE REFERENCE: 10466-14
: CURRENT APPLICATION NUMBER: US/09/907,794A
: CURRENT FILING DATE: 2001-07-17
: PRIOR APPLICATION NUMBER: PCT/US00/04414
: PRIOR FILING DATE: 2000-02-22
: PRIOR APPLICATION NUMBER: US 60/143,048
: PRIOR FILING DATE: 1999-07-07
: PRIOR APPLICATION NUMBER: US 60/145,698
: PRIOR FILING DATE: 1999-07-26
: PRIOR APPLICATION NUMBER: US 60/146,222
: PRIOR FILING DATE: 1999-07-28
: PRIOR APPLICATION NUMBER: PCT/US99/20594
: PRIOR FILING DATE: 1999-09-08
: PRIOR APPLICATION NUMBER: PCT/US99/20944
: PRIOR FILING DATE: 1999-09-13
: PRIOR APPLICATION NUMBER: PCT/US99/21030
: PRIOR FILING DATE: 1999-09-15
: PRIOR APPLICATION NUMBER: PCT/US99/21547
: PRIOR FILING DATE: 1999-09-15

```

APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/905,125A
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 185
LENGTH: 501
TYPE: PRT
ORGANISM: Homo sapiens
US-09-907-794A-185

Query Match 10.7%; Score 187.5; DB 4; Length 501;
Best Local Similarity 27.2%; Pred. No. 3.6e-11;
Matches 64; Conservative 45; Mismatches 87; Indels 39; Gaps 7;
QY 28 LSAMRLAHGCVNDTPVSTLTTPKTS-----EFENFKTKVITSKDYPLSKNFPYSLEHL 83
Db 230 LRHLKILHVSNLTVPKSNITDVAPHLTKLVHNDGTLLVLS-----LKKMM--NVAEL 283
QY 84 QTSYGLVRVDMRMLCLSLRKLDS-----ENHIKKLPA 118
Db 284 ELQCELEIRIPHAIFSLNQLQELDLKSNIRIEIISFQHLKRLTCLKLNKIVITIPP 343
QY 119 TIGDLIHLQELNLDNHLSEFVALCHSTLOKSLRSLDKSKNKIKALPVQFOQLQELKML 178
Db 344 SITHVKNLESYFSNNKLESPLVAV--FSLQK-LRCLDVSNNISMIPIEIGLQNLQHL 400
QY 179 KLDNELLQFPCKIGOLINLRLPLSARNKLPPLPSEFRNLS-LEYLDLFGNTFEQ 232
Db 401 HITGNKVDILPKQFKCIKRLNLGNCITSIPKVGQSLQTLQELKGNCLDR 455

RESULT 5
US-09-905-125A-185
Sequence 185, Application US/09905125A
Patent No. 6664376
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin J.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.

Query Match 10.7%; Score 187.5; DB 4; Length 501;
Best Local Similarity 27.2%; Pred. No. 3.6e-11;
Matches 64; Conservative 45; Mismatches 87; Indels 39; Gaps 7;
QY 28 LSAMRLAHGCVNDTPVSTLTTPKTS-----EFENFKTKVITSKDYPLSKNFPYSLEHL 83
Db 230 LRHLKILHVSNLTVPKSNITDVAPHLTKLVHNDGTLLVLS-----LKKMM--NVAEL 283
QY 84 QTSYGLVRVDMRMLCLSLRKLDS-----ENHIKKLPA 118
Db 284 ELQCELEIRIPHAIFSLNQLQELDLKSNIRIEIISFQHLKRLTCLKLNKIVITIPP 343
QY 119 TIGDLIHLQELNLDNHLSEFVALCHSTLOKSLRSLDKSKNKIKALPVQFOQLQELKML 178
Db 344 SITHVKNLESYFSNNKLESPLVAV--FSLQK-LRCLDVSNNISMIPIEIGLQNLQHL 400
QY 179 KLDNELLQFPCKIGOLINLRLPLSARNKLPPLPSEFRNLS-LEYLDLFGNTFEQ 232
Db 401 HITGNKVDILPKQFKCIKRLNLGNCITSIPKVGQSLQTLQELKGNCLDR 455

RESULT 6
US-09-902-775A-185
Sequence 185, Application US/09902775A


```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-081-149-8

Query Match
Best Local Similarity 10.2%; Score 178.5; DB 4; Length 582;
Matches 57; Conservative 26; Mismatches 63; Indels 39; Gaps 6;

QY 102 SLKLDLSHHHKKLPATIGDLHLQELNLDNHLESFVSVALCHSTQKSLRSIDLSKNK 161
Db 403 SMVELNLATNQLTKIEDVSGLVSLVLEVLISNNLLKKLPGL--GNLRK-LRELDLEENK 459
QY 162 IKALPVQFCQLOELKRLKLDNDELIOFPCKIGOLINLRFILSAARNKLPFLPSEFRNL-SL 220
Db 460 LESLPNEIAYLKDQLKLVLTNNQLTTLPRGIGHLTNLTJHLGLGENLTLHPPEIGTLENL 519
QY 221 EYLDLFGNTFEQPKVLPVVKLQAPLTLLESSARTILHNRIPIYGSHIIPFH--LCQDLDTA 278
Db 520 EELYLNDN---PNL-----HSLPPELALCSKLSIM 546
QY 279 KICVC 283
Db 547 SIENC 551

RESULT 9
US-09-081-149-7
; Sequence 7, Application US/09081149A
; Patent No. 6506889
; GENERAL INFORMATION:
; APPLICANT: Han, Min
; TITLE OF INVENTION: RAS SUPPRESSOR SUR-8
; FILE REFERENCE: UTC-02938
; CURRENT APPLICATION NUMBER: US/09/081,149A
; CURRENT FILING DATE: 1998-05-19
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURES:
; OTHER INFORMATION: Xaa at position 10 can be any amino acid.
US-09-081-149-7

Query Match
Best Local Similarity 10.2%; Score 177.5; DB 4; Length 582;
Matches 57; Conservative 26; Mismatches 63; Indels 39; Gaps 6;

QY 102 SLKLDLSHHHKKLPATIGDLHLQELNLDNHLESFVSVALCHSTQKSLRSIDLSKNK 161
Db 403 SMVELNLATNQLTKIEDVSGLVSLVLEVLISNNLLKKLPGL--GNLRK-LRELDLEENK 459
QY 162 IKALPVQFCQLOELKRLKLDNDELIOFPCKIGOLINLRFILSAARNKLPFLPSEFRNL-SL 220
Db 460 LESLPNEIAYLKDQLKLVLTNNQLTTLPRGIGHLTNLTJHLGLGENLTLHPPEIGTLENL 519
QY 221 EYLDLFGNTFEQPKVLPVVKLQAPLTLLESSARTILHNRIPIYGSHIIPFH--LCQDLDTA 278
Db 520 EELYLNDN---PNL-----HSLPPELALCSKLSIM 546
QY 279 KICVC 283
Db 547 SIENC 551

RESULT 10
US-09-331-403-2
; Sequence 2, Application US/09331403
```

```
; Patent No. 6489147
; GENERAL INFORMATION:
; APPLICANT: ALTMANN-JOHL, Regula; PHILIPPSEN, Peter; ALTHOFER,
; Henning; SEULBERGER, Harald.
; TITLE OF INVENTION: Adenylate cyclase gene, and its use
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Keil & Weinkauf
; STREET: 1101 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM AT-compatible, Pentium III processor
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect version 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/331,403
; FILING DATE: 21-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/07309
; FILING DATE: 29-DEC-1997
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1874 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-331-403-2

Query Match
Best Local Similarity 9.3%; Score 162; DB 4; Length 1874;
Matches 60; Conservative 31; Mismatches 85; Indels 38; Gaps 7;

QY 70 YPLSKNFPYSLEHLQTSYCGLVFVDMR-----MLCLKSLRKLDLSHHHKKLPATI 120
Db 762 YNKIRSLPDSMNQLOK---LAKINLSNNRIITHVNDLSKMTSLRTDLRYNRIESIKRV 817
QY 121 GDLIHLOELNLDNHLESF-----SVALCHSTLQKSLRSIDLSKNKIK 163
Db 818 P---NLQNLFLTENLTMPFDDQLMLRTLELQNPISLTLDKNDYDEHLTSLISAKLA 874
QY 164 ALPVQFC-QLOELKRLKLDNDELIOFPCKIGOLINLRFILSAARNKLPFLPSEFRNL-SLE 221
Db 875 VLPESLRLRLPRLEKLELSENSITVLPDILKHLKLVHLSVAKNKLESLEPEIAS-KNLK 934
QY 222 YLDLFGNTFEQPKVLPVVKLQAPLTLLESSARTI 255
Db 935 MLDLHCNKL---MTLPAALSTLSLTFVNISSNML 965

RESULT 11
US-09-434-408-2
; Sequence 2, Application US/09434408
; Patent No. 6440697
; GENERAL INFORMATION:
; APPLICANT: Venezia, Domenick
; APPLICANT: Grossmann, Angelika
; TITLE OF INVENTION: RING FINGER PROTEIN ZAPOP3
; FILE REFERENCE: 98-41
; CURRENT APPLICATION NUMBER: US/09/434,408
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: US 60/108,258
; EARLIER FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 723
; TYPE: PRT
```


Query Match	8.5%	Score 148.5;	DB 3;	Length 560;
Best Local Similarity	29.8%;	Fred. No. 5.6e-07;		
Matches	71;	Conservative 30;	Mismatches 82;	Indels 57; Gaps 12;
QY	29	SAMRLAHGCGNVDPVSTLTLPVKTSFEPFNKTKMTITSKKDYPLSKNFPVYSLEHLQTSYC	88	
Db	71	SGMTVLQRLMISDSHSAVAPGTFSDIKLKT-----LRLSRN---KITHLPGAL	117	
QY	89	GLVVRVDMRVLCLKSLKRLDLSHNHIKKLPATI-GDLHLQELMLNNHLESFVALCHST	147	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2004, 09:54:39 ; Search time 60 Seconds
(without alignments)
1591.684 Million cell updates/sec

Title: US-10-009-557-9

Perfect score: 1745

Sequence: 1 MYLSPVFIIQLLFIQAI.....PIISYFCSIGCVNSDMLK 338

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001s.*

5: Geneseq2002s.*

6: Geneseq2003as.*

7: Geneseq2003bs.*

8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1745	100.0	338	4	AAB27231 Human EXM
2	1660	95.1	414	4	AA04049 Human pol
3	1649	94.5	416	5	AAE17323 Human leu
4	1632	93.5	429	4	AA041835 Human pol
5	1606	92.0	308	5	ABG31598 Fruit fly
6	1315	75.4	259	5	AAE17322 Human leu
7	1137	65.2	259	4	AAU20511 Human sec
8	284.5	16.3	428	4	AB062072 Drosophil
9	218.5	12.5	389	6	ABO14780 Novel hum
10	217	12.4	1489	6	ABO14779 Novel hum
11	217	12.4	1489	6	ABU11038 Human pro
12	211.5	12.1	431	4	ABB61551 Drosophil
13	207.5	11.9	378	4	ABB65454 Drosophil
14	207	11.9	849	4	ABB64971 Drosophil
15	201.5	11.5	1060	4	ABB67403 Drosophil
16	201.5	11.5	1066	4	ABB58166 Drosophil
17	200	11.5	1052	3	AB42389 Human ORF
18	198	11.3	761	6	ABO00758 Polypepti
19	198	11.3	761	7	AB09428 Novel pro
20	197.5	11.3	474	4	AB094199 Human pro
21	197.5	11.3	546	2	AA113386 Amino aci
22	197.5	11.3	546	3	ACJ78562 Human PRO
23	197.5	11.3	546	4	AA080254 Human PRO
24	197.5	11.3	546	4	AAU29229 Human PRO
25	197.5	11.3	546	5	AAU83653 Human PRO

26	197.5	11.3	546	5	ABB04836 Human PRO
27	197.5	11.3	546	5	ABB95442 Human ang
28	197.5	11.3	546	6	ABU58605 Human PRO
29	197.5	11.3	546	6	ABU71632 Human PRO
30	197.5	11.3	546	6	ABU88153 Novel hum
31	197.5	11.3	546	6	ABU84468 Human sec
32	197.5	11.3	546	6	ABR66342 Human sec
33	197.5	11.3	546	6	ABR65732 Human sec
34	197.5	11.3	546	6	ABU99672 Human sec
35	197.5	11.3	546	6	ABU82911 Human PRO
36	197.5	11.3	546	6	ABU90032 Novel hum
37	197.5	11.3	546	6	ABU71487 Human PRO
38	197.5	11.3	546	6	ABR68281 Human sec
39	197.5	11.3	546	6	ABU96334 Novel hum
40	197.5	11.3	546	6	ABU92765 Human sec
41	197.5	11.3	546	6	ABU80800 Human PRO
42	197.5	11.3	546	6	ABO08842 Human sec
43	197.5	11.3	546	6	ABO02894 Human sec
44	197.5	11.3	546	6	ABR75048 Human sec
45	197.5	11.3	546	6	ABR94810 Human sec

ALIGNMENTS

RESULT 1

AAB27231

ID AAB27231 standard; protein; 338 AA.

XX

AC AAB27231;

XX

DT 27-MAR-2001 (first entry)

XX

DE Human EXMAD-9 SEQ ID NO: 9.

XX

KW Extracellular matrix and adhesion-associated protein; EXMAD; cancer;

KW inflammation; reproductive disorder; cardiovascular disorder;

KW immune disorder; musculoskeletal disorder; developmental disorder;

KW gastrointestinal disorder; cell proliferation disorder.

XX

OS Homo sapiens.

XX

PN WO2006068380-A2.

XX

PD 16-NOV-2000.

XX

PF 10-MAY-2000; 2000WO-US012811.

XX

PR 11-MAY-1999; 99US-0133643P.

PR

23-AUG-1999; 99US-0150409P.

XX

PA (INCY-) INCYTE GENOMICS INC.

PI

Bandman O, Hillman JL, Tang YT, Lal P, Yue H, Baughn MR, Lu DAM;

PI Azimzal Y;

XX

WPI; 2001-007395/01.

DR

N-PSDB; AAC66898.

XX

PT Isolated polynucleotide encoding extracellular matrix or adhesion-

associated protein (EXMAD) useful for diagnosing, treating, or preventing

disorders associated with expression of EXMAD such as proliferative,

immune and genetic disorders.

XX

PS Claim 1; Page 96-97; 129pp; English.

XX

The present invention provides the protein and coding sequences for 25 novel extracellular matrix and adhesion-associated proteins (EXMADs). These are designated EXMAD-1, EXMAD-2, EXMAD-3, EXMAD-4, EXMAD-5, EXMAD-6, EXMAD-7, EXMAD-8, EXMAD-9, EXMAD-10, EXMAD-11, EXMAD-12, EXMAD-13, EXMAD-14, EXMAD-15, EXMAD-16, EXMAD-17, EXMAD-18, EXMAD-19, EXMAD-20, EXMAD-21, EXMAD-22, EXMAD-23, EXMAD-24 and EXMAD-25. They are useful in the prevention and treatment of cancers, cell proliferation,

CC cardiovascular, reproductive, immune, musculoskeletal, developmental and
 CC gastrointestinal disorders and inflammation
 XX
 XX Sequence 338 AA;
 Query Match 100.0%; Score 1745; DB 4; Length 338;
 Best Local Similarity 100.0%; Pred. No. 5.8e-169;
 Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MYVLSVPEFIIQLLFQIAQLSSLSKGLFSAMRLAHRCNVDTVPVSTLTPVKTSEFENFKT 60
 Db |||||
 QY 1 MYVLSVPEFIIQLLFQIAQLSSLSKGLFSAMRLAHRCNVDTVPVSTLTPVKTSEFENFKT 60
 Db |||||
 QY 61 KMVITSKKYPLSKNFPYSLHQLTSYCGLVRRVDMRLCLKSLRKLDSHNHKKLPAT 120
 Db |||||
 QY 61 KMVITSKKYPLSKNFPYSLHQLTSYCGLVRRVDMRLCLKSLRKLDSHNHKKLPAT 120
 Db |||||
 QY 121 GOLIHLCLELNDNHLESVALCHSTLQKSLRSLDLSSKNIKALPVQFCQLQELKNKL 180
 Db |||||
 QY 121 GOLIHLCLELNDNHLESVALCHSTLQKSLRSLDLSSKNIKALPVQFCQLQELKNKL 180
 Db |||||
 QY 181 DONELTQFPCKIGQLINLFLSARNKLPLPSEFENLSLEYLDLFGNTPEQKVLPLVIK 240
 Db |||||
 QY 181 DONELTQFPCKIGQLINLFLSARNKLPLPSEFENLSLEYLDLFGNTPEQKVLPLVIK 240
 Db |||||
 QY 241 LQAPLTLESSARTILHNRIPIYGSHIIPFHLQDLDLTKICVCGFCLNSFIQGTITMNL 300
 Db |||||
 QY 241 LQAPLTLESSARTILHNRIPIYGSHIIPFHLQDLDLTKICVCGFCLNSFIQGTITMNL 300
 Db |||||
 QY 301 HSAHTVVLVDNLGGTEAPIISYFCSLGCYVNSDDMLK 338
 Db |||||
 QY 301 HSAHTVVLVDNLGGTEAPIISYFCSLGCYVNSDDMLK 338
 Db |||||

RESULT 2
 ID AAM40049 standard; protein; 414 AA.
 XX AAM40049;
 AC AAM40049;
 XX AAM40049;
 DT 22-OCT-2001 (first entry)
 DE Human polypeptide SEQ ID NO 3194.
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX Homo sapiens.
 OS Homo sapiens.
 XX WO200153312-A1.
 PN 26-JUL-2001.
 PD 26-JUL-2001.
 XX 26-DEC-2000; 2000WO-US034263.
 XX 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

PI Zhou P, Goodrich R, Drmanac RT;
 XX MPI; 2003-442253/47.
 DR N-PSDB; AAI59205.
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX Example 4; SEQ ID NO 3194; 10078pp; English.
 SS The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX Sequence 414 AA;

Query Match 95.1%; Score 1660; DB 4; Length 414;
 Best Local Similarity 99.7%; Pred. No. 3.6e-160;
 Matches 320; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 18 QATSSSLKGLFSAMRLAHRCNVDTVPVSTLTPVKTSEFENFKTMMVITSKKYPLSKNFP 77
 Db |||||
 QY 94 KATSSSLKGLFSAMRLAHRCNVDTVPVSTLTPVKTSEFENFKTMMVITSKKYPLSKNFP 153
 Db |||||
 QY 78 YSLEHLQTSYCGLVRRVDMRLCLKSLRKLDSHNHKKLPATIGDLIHQELNNDNHLE 137
 Db |||||
 QY 154 YSLEHLQTSYCGLVRRVDMRLCLKSLRKLDSHNHKKLPATIGDLIHQELNNDNHLE 213
 Db |||||
 QY 138 SFSVALCHSTLQKSLRSLDLSSKNIKALPVQFCQLQELKNKLDDNELIQFCKIGQLIN 197
 Db |||||
 QY 214 SFSVALCHSTLQKSLRSLDLSSKNIKALPVQFCQLQELKNKLDDNELIQFCKIGQLIN 273
 Db |||||
 QY 198 LAFLSARNKLPLPSEFENLSLEYLDLFGNTPEQKVLPLVIKQAPLTLESSARTILH 257
 Db |||||
 QY 274 LAFLSARNKLPLPSEFENLSLEYLDLFGNTPEQKVLPLVIKQAPLTLESSARTILH 333
 Db |||||
 QY 258 NRIPYGSHIIPFHLQDLDLTKICVCGFCLNSFIQGTITMNLHSAHTVVLVDNLGGTE 317
 Db |||||
 QY 334 NRIPYGSHIIPFHLQDLDLTKICVCGFCLNSFIQGTITMNLHSAHTVVLVDNLGGTE 393
 Db |||||
 QY 318 APIISYFCSLGCYVNSDDMLK 338
 Db |||||
 QY 394 APIISYFCSLGCYVNSDDMLK 414

RESULT 3
 ID AAE17323 standard; protein; 416 AA.
 XX AAE17323;
 AC AAE17323;
 XX AAE17323;
 DT 18-APR-2002 (first entry)
 DE Human leucine rich repeat: protein, sbg442358PROA #2.
 XX Human; therapy; wound healing disorder; vaccine; cancer; infection;
 KW autoimmune disorder; haematopoietic disorder; inflammation; arthritis;
 KW Parkinson's disease; Huntington's chorea; schizophrenia; arrhythmic;
 KW multiple sclerosis; Alzheimer's disease; atherosclerosis; brain disorder;
 KW ischaemia; stroke; AIDS; bone disease; attherosclerosis; brain disorder;
 KW depression; cardiovascular disease; myocardial infarction; renal failure;

respiratory disease; liver disorder; Fanconi's syndrome; spleen disorder; type II diabetes mellitus; skeletal muscle disorder; immunosuppressive; hyperlipemia; renal disease; hypoglycaemia; gastrointestinal disease; neutropenic; cirrhosis; Hodgkin's disease; neuroleptic; anti-inflammatory; haemostatic; vulnery; anticonvulsant; antirheumatic; neuroprotective; nephrotropic; hypotensive; vasotropic; cytostatic; cerebroprotective; allergy; leucine rich repeat protein.

Homo sapiens.

WO2001198342-A1.

27-DEC-2001.

22-JUN-2001; 2001WO-US019929.

22-JUN-2000; 2003US-0213156P.

22-JUN-2000; 2003US-0213161P.

(SMIK) SMITHKLINE BEECHAM CORP.

(SMIK) SMITHKLINE BEECHAM PLC.

(GLAX) GLAXO GROUP LTD.

Aagarwal P, Cogswell JP, Kabric KS, Lai Y, Martensen SA; Murdock PR, Smith RF, Strum JC, Xiang Z, Xie Q, Rizni SK; WPI; 2002-139783/18.

N-PSDB; AAD27818.

Novel secreted and membrane-associated polypeptides and polynucleotides useful for preventing, ameliorating or correcting dysfunction or disease including diabetes, cancer, hypertension and growth abnormalities.

Claim 1; Page 134-135; 138pp; English.

The invention relates to secreted and membrane-associated polypeptides and polynucleotides. The sequences of the invention are useful in diagnostic assays for detecting diseases associated with inappropriate activity or levels of these polynucleotides, and in identifying their agonists and antagonists that are potentially useful in therapy. The sequences of the invention are useful as vaccines for inducing immunological response. The sequences of the invention are useful for treating cancers, infections, autoimmune disorders, haematopoietic disorders, wound healing disorders, cholesteryl ester storage disease, inflammation, congenital muscular dystrophy, junctional epidermolysis bullosa, Parkinson's disease, Huntington's chorea, multiple sclerosis, viral and bacterial infections, Alzheimer's disease, asthma, arthritis, allergies, schizophrenia, sbg44245PBOA-associated disorders, septicæmia, psoriasis, inflammatory bowel disease, transplant rejection, graft verse host disease, ischaemia, stroke, acute respiratory disease syndrome, restenosis, brain injury, AIDS, bone diseases, atherosclerosis, brain disorders including paraspranuclear paley, myotonic dystrophy, depression, anxiety disorders and sleep disorders, cardiovascular diseases including congestive heart failure and myocardial infarction, respiratory diseases including chronic obstructive pulmonary disease, acute bronchitis and adult respiratory distress syndrome, liver disorders including hypercholesterolaemia, hypertriglyceridaemia, cirrhosis, viral and non-viral hepatitis, type II diabetes mellitus, renal disease including acute and chronic renal failure, glomerulonephritis, Fanconi's syndrome, cystinuria, skeletal muscle disorders including hypoglycaemia and tendinitis, gastrointestinal diseases including intestinal obstruction and tropical sprue, spleen disorders including hypersplenism, Hodgkin's disease and malignant lymphoma, testicular cancer, male reproductive diseases including low testosterone and male infertility. The present sequence is human leucine rich repeat protein

Db	94	KAISSSLKGFLSAMRLAHRGCNVDPVSTLTPEVKTSPENFKTKMVIITSKDYPLSKNFP	153
Qy	78	YLSLHLQTSYGLVRDVMCLKSLRKLDLSHNEIKKLPATIGDLIHLQELNNDNHLE	137
Db	154	YLSLHLQTSYGLVRDVMCLKSLRKLDLSHNEIKKLPATIGDLIHLQELNNDNHLE	213
Qy	138	SFSVALCHSTLQKSLRSLDLSKNKIKALPVQFCQLQELNKLKLDNELIQFPCKISQJIN	197
Db	214	SFSVALCHSTLQKSLRSLDLSKNKIKALPVQFCQLQELNKLKLDNELIQFPCKISQJIN	273
Qy	198	LRFLSAARNKLPFLPSEFRNLSLEYLDLFGNTFEQPKVLPVIKQAPLTILESSARTILH	257
Db	274	LRFLSAARNKLPFLPSEFRNLSLEYLDLFGNTFEQPKVLPVIKQAPLTILESSARTILH	333
Qy	258	--NRIPYGSIIIPHLQCDLDTAKICVCGRFCLNSFIQGTITMKLHSAHTVWLVDNLGG	315
Db	334	NRNRIPYGSHIIPHLQCDLDTAKICVCGRFCLNSFIQGTITMKLHSAHTVWLVDNLGG	393
Qy	316	TEAPIISYFCSLGCVNNSDMLK	338
Db	394	TEAPIISYFCSLGCVNNSDMLK	416
RESULT 4			
AAM41835			
ID	AAM41835 standard; protein; 429 AA.		
XX			
AC	AAM41835;		
XX			
DT	22-OCT-2001 {first entry}		
XX			
DE	Human polypeptide SEQ ID NO 6766.		
XX			
KW	Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;		
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;		
leukaemia.			
XX			
OS	Homo sapiens.		
XX			
PN	WO200153312-A1.		
XX			
PD	26-JUL-2001.		
XX			
PF	26-DEC-2000; 2000WO-US034263.		
XX			
PR	23-DEC-1999; 99US-00471275.		
PR	21-JAN-2000; 2000US-00488725.		
PR	25-APR-2000; 2000US-00552317.		
PR	20-JUN-2000; 2000US-00598042.		
PR	19-JUL-2000; 2000US-00620312.		
PR	03-AUG-2000; 2000US-00653450.		
PR	14-SEP-2000; 2000US-00662191.		
PR	19-OCT-2000; 2000US-00693036.		
PR	29-NOV-2000; 2000US-00727344.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
PI	Wang J, Wang Z, Wehrman I, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;		
PI	Zhou P, Goodrich R, Drmanac RT;		
XX			
XX	WPI; 2001-442253/47.		
DR	N-PSDB; AAI60991.		
XX			
PT	Novel nucleic acids and polypeptides, useful for treating disorders such		
PT	as central nervous system injuries.		
XX			
PS	Example 2; SEQ ID NO 6766; 10078pp; English.		
XX			

CC The invention relates to human nucleic acids (AA157798-AA161369) and the
 CC encoded polypeptides (AA157798-AA161369) with neurotrophic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 SQ Sequence 429 AA;

Query Match 93.5%; Score 1632; DB 4; Length 429;
 Best Local Similarity 98.7%; Pred. No. 2.7e-157;
 Matches 315; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 18 QATSSSLKGLSAMRLAHRGNDVTPVSTLTPVKTSSEFENFKTKWITSKDYPLSKNFP 77
 Db 111 KANSSSLKGLSAMRLAHRGNDVTPVSTLTPVKTSSEFENFKTKWITSKDYPLSKNFP 172
 QY 78 YSLEHLQTSYCGLVVRVDMRMCLSKSLKLDLSHNHKKLPATIGDLHLQELNLDNHLE 137
 Db 171 YSLEHLQTSYCGLVVRVDMRMCLSKSLKLDLSHNHKKLPATIGDLHLQELNLDNHLE 230
 QY 138 SFSVALCHSTLQKSLSLKSKNKKALPVQFCOLOELKNKLDNDNELIOFFCKIGQLIN 197
 Db 231 SFSVALCHSTLQKSLSLKSKNKKALPVQFCOLOELKNKLDNDNELIOFFCKIGQLIN 290
 QY 198 LRFLSAARNKLPPLPSEFRNLSEYLDLFGNTFEPQKLVPIKQAPLTLLSSARTILH 257
 Db 291 LRFLSAARNKLPPLPSEFRNLSEYLDLFGNTFEPQKLVPIKQAPLTLLSSARTILH 350
 QY 258 NRIPYSHIIPFHLQDLDLAKICVGRFCNLSFIQGTITMNLHSAHTVVLVDNLGGTE 317
 Db 351 NRIPYSHIIPFHLQDLDLAKICVGRFCNLSFIQGTITMNLHSAHTVVLVDNLGGTE 410
 QY 318 APIISYFCSLGCVNNSDM 336
 Db 411 APIISYFCSLGCVNNSDI 429

RESULT 5
 ABG31598
 ID ABG31598 standard; protein; 308 AA.
 AC ABG31598;

DT 05-NOV-2002 (first entry)

DE Fruit fly LRR47 polypeptide 47-33.88.

XX Fruit fly; LRR47 polypeptide 47-33.88; embryonic development deformity;
 KW tumour; diabetes; menstrual disorder; peptide ulcer; arrhythmia; anaemia;
 KW epilepsy.

OS Drosophila sp.

PN CN1341640-A.

PD 27-MAR-2002.

XX 05-SEP-2000; 2000CN-00125025.

PF 05-SEP-2000; 2000CN-00125025.

PR 05-SEP-2000; 2000CN-00125025.

XX (SHAN-) SHANGHAI BIODOOR GENE DEV CO LTD.

PI Mao Y, Xie Y;

XX WPI; 2002-520716/56.

DR N-PSDB; ABK90911.

XX A fruit fly LRR47 polypeptide 47-33.88, useful for curing e.g. tumors and
 PT diabetes.

XX Claim 1; Page 26-27 (Disclosure); 33pp; Chinese.

XX The present invention relates to a new fruit fly LRR47 polypeptide 47-
 CC 33.88. The polypeptide is useful for curing several diseases, such as
 CC embryonic development deformity, tumour, diabetes, menstrual disorder,
 CC peptide ulcer, arrhythmia, anaemia and epilepsy. The present amino acid
 CC sequence represents the fruit fly LRR47 polypeptide 47-33.88 of the
 CC invention

XX Sequence 308 AA;

Query Match 92.0%; Score 1606; DB 5; Length 308;
 Best Local Similarity 100.0%; Pred. No. 7.5e-155;
 Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 31 MRLAHRGNDVTPVSTLTPVKTSSEFENFKTKWITSKDYPLSKNFPYSLEHLOTSYGL 90
 Db 1 MRLAHRGNDVTPVSTLTPVKTSSEFENFKTKWITSKDYPLSKNFPYSLEHLOTSYGL 60
 QY 91 VRVDMRMCLSKSLKLDLSHNHKKLPATIGDLHLQELNLDNHLESFVALCHSTLQK 150
 Db 61 VRVDMRMCLSKSLKLDLSHNHKKLPATIGDLHLQELNLDNHLESFVALCHSTLQK 120
 QY 151 SLRSLDLSKNNKIKALPVQFCOLOELKNKLDNDNELIOFFCKIGQLINLRFLSAARNKLP 210
 Db 121 SLRSLDLSKNNKIKALPVQFCOLOELKNKLDNDNELIOFFCKIGQLINLRFLSAARNKLP 180
 QY 211 LPSEFRNLSEYLDLFGNTFEPQKLVPIKQAPLTLLSSARTILHNRIPYSGHIIPFH 270
 Db 181 LPSEFRNLSEYLDLFGNTFEPQKLVPIKQAPLTLLSSARTILHNRIPYSGHIIPFH 240
 QY 271 LCQDLDLAKICVGRFCNLSFIQGTITMNLHSAHTVVLVDNLGGTEAPIISYFCSLGCV 330
 Db 241 LCQDLDLAKICVGRFCNLSFIQGTITMNLHSAHTVVLVDNLGGTEAPIISYFCSLGCV 300
 QY 331 VNSSDMLK 338
 Db 301 VNSSDMLK 308

RESULT 6
 AAE17322
 ID AAE17322 standard; protein; 259 AA.
 AC AAE17322;

DT 19-APR-2002 (first entry)

DE Human leucine rich repeat protein, sbg442358PROA #1.

XX Human; therapy; wound healing disorder; vaccine; cancer; infection;
 KW autoimmune disorder; haematopoietic disorder; inflammation; arthritis;
 KW Parkinson's disease; Huntington's chorea; schizophrenia; antiarrhythmic;
 KW multiple sclerosis; Alzheimer's disease; analgesic; cardiac; asthma;
 KW ischaemia; stroke; AIDS; bone disease; atherosclerosis; brain disorder;
 KW depression; cardiovascular disease; myocardial infarction; renal failure;
 KW respiratory disease; liver disorder; Fanconi's syndrome; spleen disorder;
 KW type II diabetes mellitus; skeletal muscle disorder; immunosuppressive;
 KW hyperlipidemia; renal disease; hypoglycaemia; gastrointestinal disease;
 KW neoplastic; cirrhosis; Hodgkin's disease; neuroleptic; antiinflammatory;
 KW haemostatic; vulnery; anticonvulsant; antirheumatic; neuroprotective;
 KW nephrotropic; hypotensive; vasotropic; cyostatic; cerebroprotective;
 KW allergy; leucine rich repeat protein.

XX Homo sapiens.

XX WO200198342-A1.
 XX 27-DEC-2001.
 XX 22-JUN-2000; 2001WO-US019929.
 XX 22-JUN-2000; 2000US-0213156P.
 XX 22-JUN-2000; 2000US-0213161P.
 XX (SMLK) SMITHKLINE BEECHAM CORP.
 XX (SMLK) SMITHKLINE BEECHAM PCC.
 XX (GLAX) GLAXO GROUP LTD.
 XX Agarwal P, Cogswell JP, Kabnic KS, Lai Y, Martensen SA;
 XX Murdock PR, Smith RF, Strum JC, Xiang Z, Xie Q, Rizni SK;
 XX WPI: 2002-139783/18.
 XX N-PSDB; AAD27817.
 XX Novel secreted and membrane-associated polypeptides and polynucleotides
 XX useful for preventing, ameliorating or correcting dysfunction or disease
 XX including diabetes, cancer, hypertension and growth abnormalities.
 XX Claim 1: Page 134; 138pp; English.
 XX The invention relates to secreted and membrane-associated polypeptides
 XX and polynucleotides. The sequences of the invention are useful in
 XX diagnostic assays for detecting diseases associated with inappropriate
 XX activity or levels of these polynucleotides, and in identifying their
 XX agonists and antagonists that are potentially useful in therapy. The
 XX sequences of the invention are useful as vaccines for inducing
 XX immunological response. The sequences of the invention are useful for
 XX treating cancers, infections, autoimmune disorders, haematopoietic
 XX disorders, wound healing disorders, cholesterol ester storage disease,
 XX inflammation, congenital muscular dystrophy, junctional epidermolysis
 XX bullosa, Parkinson's disease, Huntington's chorea, multiple sclerosis,
 XX viral and bacterial infections, Alzheimer's disease, asthma, arthritis,
 XX allergies, schizophrenia, sbg4445PR0a-associated disorders,
 XX septicemia, psoriasis, inflammatory bowel disease, transplant rejection,
 XX graft versus host disease, ischaemia, stroke, acute respiratory disease,
 XX syndrome, restenosis, ischaemia, stroke, acute respiratory disease,
 XX brain disorders including paraneuronal palsy, myotonic dystrophy,
 XX depression, anxiety disorders and sleep disorders, cardiovascular
 XX diseases including congestive heart failure and myocardial infarction,
 XX respiratory diseases including chronic obstructive pulmonary disease,
 XX acute bronchitis and adult respiratory distress syndrome, liver disorders
 XX including hypercholesterolaemia, hypertriglyceridaemia, cirrhosis, viral
 XX and non-viral hepatitis, type II diabetes mellitus, renal disease
 XX including acute and chronic renal failure, glomerulonephritis, Fanconi's
 XX syndrome, cystinuria, skeletal muscle disorders including hypoglycaemia
 XX and tendinitis, gastrointestinal diseases including intestinal
 XX obstruction and tropical sprue, spleen disorders including hypersplenism,
 XX Hodgkin's disease and malignant lymphoma, testicular cancer, male
 XX reproductive diseases including low testosterone and male infertility.
 XX The present sequence is human leucine rich repeat protein

Sequence 259 AA;

Query Match 75.4%; Score 1315; DB 5; Length 259;
 Best Local Similarity 100.0%; Pred. No. 2.7e-125;
 Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYVLSPEFIIQLLFIQAISSSLKGFSLSMRLAHRGNCVDPVSTLTPVKTSSEFENFT 60
 Db 1 MYVLSPEFIIQLLFIQAISSSLKGFSLSMRLAHRGNCVDPVSTLTPVKTSSEFENFT 60
 Qy 61 KMWITSKXDYPLSKNFPYSLEHQTSTYGLVRVDMRLCLSLRKLDLSHNHKKLPATI 120
 Db 61 KMWITSKXDYPLSKNFPYSLEHQTSTYGLVRVDMRLCLSLRKLDLSHNHKKLPATI 120
 Qy 121 GDLHLQELNLDNHLESFVALCHSTLQKSLRSLDLKSNKIKALPVQFCOLQELKNKL 180

Db 121 GDLHLQELNLDNHLESFVALCHSTLQKSLRSLDLKSNKIKALPVQFCOLQELKNKL 180
 Qy 181 DONELIQPCKIGQLINLRFLSARNKLPPLPSEFRNLSLEYLDLFGNTPEQPKVLPIK 240
 Db 181 DONELIQPCKIGQLINLRFLSARNKLPPLPSEFRNLSLEYLDLFGNTPEQPKVLPIK 240
 Qy 241 LQAPLTLESSARTILHNR 259
 Db 241 LQAPLTLESSARTILHNR 259
 RESULT 7
 AAU20511
 ID AAU20511 standard; protein; 259 AA.
 AC AAU20511;
 XX 06-DEC-2001 (first entry)
 DT XX
 XX Human secreted protein, Seq ID No 503.
 DE XX
 XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
 KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;
 KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
 KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
 KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
 KW Gaucher's disease; neurological disease; cerebrovascular disorder;
 KW thrombosis; wound healing.
 XX XX
 OS Homo sapiens.
 PN WO200155326-A2.
 XX 02-AUG-2001.
 PD XX
 PF 17-JAN-2001; 2001WO-US001347.
 XX 31-JAN-2000; 2000US-017906SP.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 11-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 14-JUL-2000; 2000US-0217496P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-022668-P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 03-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239337P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246533P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.

PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 06-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254037P.
 PR 05-JAN-2001; 2001US-02559678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-451931/48.
 DR N-PSDB; AAS33220.
 XX
 FT
 PT
 XX
 PS
 XX

New nucleic acids and polypeptides, useful for diagnosing, preventing or treating medical conditions.

Claim 11; SEQ ID NO 503; 753pp; English.

The invention relates to novel isolated nucleic acid molecules (I) encoding human secreted proteins (II). (I) and (II) are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate expression of secreted proteins. (I) and complementary sequences may also be used as DNA probes in diagnostic assays (e.g. polymerase chain reactions (PCR)) to detect and quantitate the presence of similar nucleic acid sequences in samples, and so which patients may be in need of restorative therapy. (II) may also be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of the expression and activity of the secreted proteins. The anti-(II) antibodies and antagonists may also be used to down regulate expression and activity of (II). The anti-(II) antibodies may also be used as diagnostic agents for detecting the presence of (II) in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). The disorders include for example: immune/autoimmune diseases (e.g. HIV (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis and multiple sclerosis), cancers and hyperproliferative disorders (e.g. melanomas, neoplasms of the breast or liver, Sezary syndrome and Gaucher's disease), neurological diseases (e.g. Alzheimer's disease, Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and thrombosis), infections caused by bacteria, viruses and fungi and ocular disorders (e.g. corneal infections). (I) and (II), agonists, antagonists and antibodies can also be used to promote wound healing, maintain organs before transplantation, and support cell culture of primary tissues.

Query Match 65.2%; Score 1137; DB 4; Length 259;
 Best Local Similarity 98.7%; Pred. No. 3.8e-107;
 Matches 222; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 18 QAISSSLKGLSAMLRAHRCNVDPVSTLTPVKCTSEPFNTKMTWTSKKDYPLSKNFP 77

Db 33 KAMSSSLKGFUSARLHARGCNVDTPTVSTLTPVKTSFENPKTKRWITSKDYPLSKNFP 92
 QY 78 YSLEHLQTSYCGLVVRVDMRMLCLKSLRKLDSLHNHKKLPATIGDLHLQELNLDNHL 137
 Db 93 YSLEHLQTSYCGLVVRVDMRMLCLKSLRKLDSLHNHKKLPATIGDLHLQELNLDNHL 152
 QY 138 SFSVALCHSTLQKSLRSLDLSKNKIKALPVQFCQLQELKNLKLDDNELIQPCKIGQJIN 197
 Db 153 SFSVALCHSTLQKSLRSLDLSKNKIKALPVQFCQLQELKNLKLDDNELIQPCKIGQJIN 212
 QY 198 LRFLSAARNKLPFLPSEFRNLSLEYLDLFGNTFPQPKVLPIKIQ 242
 Db 213 LRFLSAARNKLPFLPSEFRNLSLEYLDLFGNTFPQPKVLPIKIQ 257

RESULT 8

AB62072
 ID AB62072 standard; protein; 428 AA.
 XX AC AB62072;
 XX DT 26-MAR-2002 (first entry)
 XX DE Drosophila melanogaster polypeptide SEQ ID NO 13008.
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 XX KW pharmaceutical.
 XX OS Drosophila melanogaster.
 XX ZN WO200171042-A2.

XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US039231.
 XX 23-MAR-2000; 2000US-0191637P.
 XX 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 XX Venter JC; Adams M, Li PWD, Myers EW;
 XX WPI; 2001-856863/75.
 XX N-PSDB; ABL06175.

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.

XX Disclosure; SEQ ID NO 13008; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 428 AA;

Query Match 16.3%; Score 284.5; DB 4; Length 428;

Best Local Similarity 28.7%; Pred. No. 6.4e-20;
 Matches 96; Conservative 53; Mismatches 135; Indels 51; Gaps 10;

QY 15 LFIQAISSSLKGFLSAMRLAHRG-----CNVDTPTVSTLTPVKTSFENPKTKRWITSKDK 69
 Db 91 LLIKCDPTQLKGFSLQTLKMGDGDALNRLNINAATAIPQKAP-----QVRNVIKRSK 146

QY 70 YPLSKNFPYSLEH-QTSYCGLVVRVDMRMLCLKSLRKLDSLHNHKKLPATIGDLHLQEL 129
 Db 147 YPI-KGFRTILKSLTINNSQLVLSFEICTLRLNLTKLDVSGNKLTKIPSELRL-PLTSL 204
 QY 130 NLNDNHLESFS-----VALCHSTLQKSLRSLDLSKNKIKALPVQFCQLQELKNLKLDD 182
 Db 205 HLGNLLGTQNDWCWLRGTCLC-----QSLGELDLSGNGLTYFPPLPKFESLSVLNN 259
 QY 183 NELIQPCKIGQJINLRFLSAARNKLPFLPSEFRNLSLEYLDLFGNTFE-----QPKV 235
 Db 260 NLLSRLPFAIRRMKALRKLXVCSELESLSAVEDLRIDLDVGMGNCFKBNADAQOMY 319
 QY 236 LPVIKLOAPLTLESSARTILENRPVGSHPHLCQDLDTAKICVGRFCL----- 288
 Db 320 LQKAASNSQPLMLLGRADVDMYMLPISAGSIPAVLIDIREAPCPGELCYAORKEDL 379
 QY 289 -----NSFIQGTMMNL-----HSVHTVVVLVDN 312
 Db 380 QRVVQPKFI---TVKNLTYSREHQAIVADVVLCD 411

RESULT 9

AB014780
 ID AB014780 standard; protein; 389 AA.

XX AC AB014780;

XX DT 25-AUG-2003 (first entry)

XX DE Novel human protein #153.

XX KW Human; NOV; gene therapy; endocrine related disease; diabetes; metabolism-related disease; obesity; central nervous system disorder; Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis; schizophrenia; depression; autoimmune disorder; inflammatory disorder; psoriasis; allergy; lupus erythematosus; asthma; cancer;
 KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
 KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
 KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
 KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;
 KW stroke; infection.

XX OS Homo sapiens.

XX WO2003023002-A2.

XX 26-MAR-2003.

XX 09-SEP-2002; 2002WO-US028539.

XX 07-SEP-2001; 2001US-0318120P.

XX 10-SEP-2001; 2001US-0318130P.

XX 17-SEP-2001; 2001US-0322636P.

XX 17-SEP-2001; 2001US-0322781P.

XX 17-SEP-2001; 2001US-0322816P.

XX 19-SEP-2001; 2001US-0322817P.

XX 20-SEP-2001; 2001US-0323519P.

XX 20-SEP-2001; 2001US-0323631P.

XX 25-SEP-2001; 2001US-0324969P.

XX 26-SEP-2001; 2001US-0325091P.

XX 26-SEP-2001; 2001US-0324990P.

XX 17-APR-2002; 2002US-0373212P.

XX 06-SEP-2002; 2002US-06236177.

XX {CURA-} CURAGEN CORP.

PI Spytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;
 PI Gerlach VL, Vernet CAM, Bergts C, Rothenberg ME, Guo X;
 PI Shimkova RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;
 PI Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;

[illegible]

QY 75 NFPYSLEHLQTSYCGLVVRVDMRLCLSLRKLDSLHNHKKLPATIGDLHLQELNLDN 134
 Db 43 NFERTLEELYLDANQIEELPKQLENQALRKLSTPDNDLSNLTPTTASLVNKLKELDISKN 102
 QY 135 HLESF--SVALCHSTLOKSLRSLDLKKNKIKALPVQFCOLQELKLNKLDNELLQFPCKI 192
 Db 103 GVQEPFENIKCC-----KCLTITIEASVNPISKLPGFTQLNLTLQVLYNDAFLFELPANF 157
 QY 193 GOLINRLFLSAARNKLPFLPSEFNLSLEYLDLFGNTF-EQPKVLPVTKLQAPLTLLSS 251
 Db 158 GRUVKRLILELRENHLKTLPMWKLQAQLERLDLGNNEFSELPVLDQLQNLREL-WMDNN 216
 QY 252 ARTILHNRIPIVGSIIIFHLQCD-LDTAKICVCGRFCLNSFIQGTITMNLHLSVAHTVYV 310
 Db 217 ALQVLPGSIGKMLVLDMSKNRIETVDMDISGCEALDLSSNMLQ-----QLP 268
 QY 311 ONLGG 315
 Db 269 DSIIG 273

RESULT 11
 ABU11038
 ID ABU11038 standard; protein; 1489 AA.
 AC ABU11038;
 DT 05-FEB-2003 (first entry)
 XX Human protein NOV7.
 XX Human, NOVX; adrenoleukodystrophy; haemophilia; stroke; VHL;
 KW congenital adrenal hyperplasia; haemophilia; hypercoagulation;
 KW idiopathic thrombocytopenic purpura; autoimmune disease; allergy;
 KW immunodeficiencies; transplantation; Von Hippel-Lindau syndrome;
 KW Alzheimer's disease; tubular sclerosis; Parkinson's disease; epilepsy;
 KW Huntington's disease; cerebral palsy; Lesch-Nyhan syndrome; pain;
 KW multiple sclerosis; ataxia-telangiectasia; leukodystrophy; anxiety;
 KW behavioural disorder; addiction; neuroprotection; diabetes; ARDS;
 KW renal artery stenosis; interstitial nephritis; glomerulonephritis;
 KW polycystic kidney disease; systemic lupus erythematosus; IgA;
 KW renal tubular acidosis; immunoglobulin A nephropathy; hypercalcaemia;
 KW cirrhosis; transplantation; asthma; emphysema; scleroderma; GVHD;
 KW adult respiratory distress syndrome; graft versus host disease;
 KW lymphedema; fertility; pancreatitis; obesity; haemophilia; ulcer;
 KW anaemia; cancer; trauma; regeneration; infection.
 XX Homo sapiens.
 OS
 XX
 XX WC200281629-AA.
 XX
 PD 17-OCT-2002.
 XX
 XX
 XX 03-APR-2002; 2002W0-US010522.
 XX
 XX 03-APR-2001; 2001US-0281086P.
 XX 03-APR-2001; 2001US-0281136P.
 XX 05-APR-2001; 2001US-0281863P.
 XX 06-APR-2001; 2001US-0281906P.
 XX 10-APR-2001; 2001US-0282020P.
 XX 12-APR-2001; 2001US-0282934P.
 XX 12-APR-2001; 2001US-0283512P.
 XX 19-APR-2001; 2001US-0285325P.
 XX 23-APR-2001; 2001US-0285890P.
 XX 24-APR-2001; 2001US-0286068P.
 XX 25-APR-2001; 2001US-0286292P.
 XX 27-APR-2001; 2001US-0287213P.
 XX 02-MAY-2001; 2001US-0288257P.
 XX 12-MAY-2001; 2001US-0291134P.
 XX 17-MAY-2001; 2001US-0291725P.
 XX 31-MAY-2001; 2001US-0294771P.
 XX 08-JUN-2001; 2001US-0296965P.
 XX 18-JUN-2001; 2001US-0299128P.

12-JUL-2001; 2001US-0335063P.
 14-NOV-2001; 2001US-0332780P.
 04-JAN-2002; 2002US-0345221P.
 02-APR-2002; 2002US-00345221.
 XX (CURA-) CURAGEN CORP.
 PA Spytke KA, Li L, Edinger SR, Ellerman K, Stone DJ, Malyankar UM;
 PI Shimkets RA, Guo X, Anderson DW, Patuturajac M, Berghs C, Gerlach V;
 PI Taupier RJ, Pena CBA, Padigaru M, Liu Y, Burgess CE, Miller CE;
 PI Gusev VV, Xekuda R, Gorman L, Zerhusen BD, Baumgartner JC;
 PI Tcherenev VT, Vernet CAM, Smithson G, Heyes ME, Shenoy SG, Liu X;
 PI Gangolli EA;
 DR MPI: 2003-046863/04.
 DR N-PSDB; ABX17516.
 XX
 PT New polypeptides, designated NOVX polypeptides, useful for treating
 PT hemophilia, idiopathic thrombocytopenic purpura, autoimmune disease,
 PT allergies, transplantation, Alzheimer's disease and stroke.

Claim 1; Page 102-103; 320pp; English.

The invention relates to an isolated NOVX polypeptide selected from NOV1-NOV27 polypeptides, a mature form of NOVX, a variant of NOVX or a fragment of NOVX. Also included are determining the presence or amount of NOVX in a sample (by using an antibody that immunospecifically bind to the polypeptide), determining the presence of or predisposition to disease associated with altered levels of NOVX in a first mammalian subject, identifying a potential therapeutic agent for use in the treatment of pathology related to aberrant expression of physiological interactions of NOVX, screening for a modulator of activity or of latency or predisposition to a pathology associated with NOVX, the nucleic acid encoding NOVX, vectors and host cells. NOVX is useful for identifying an agent (a cellular receptor or downstream effector) that binds to NOVX. NOVX and NOVX nucleic acids are useful for treating or preventing NOVX-associated disorders in humans, and in the manufacture of a medicament for treating a NOVX related disease human disease e.g. adrenoleukodystrophy, congenital adrenal hyperplasia, haemophilia, hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, stroke, tubular sclerosis, Parkinson's disease, Huntington's disease, cerebral palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia-telangiectasia, leukodystrophies, behavioural disorders, addiction, anxiety, pain, neuroprotection, diabetes, renal artery stenosis, interstitial nephritis, glomerulonephritis, polycystic kidney disease, systemic lupus erythematosus, renal tubular acidosis, immunoglobulin (Ig) A nephropathy, hypercalcaemia, cirrhosis, transplantation, asthma, emphysema, scleroderma, adult respiratory distress syndrome (ARDS), graft versus host disease (GVHD), lymphedema, fertility, pancreatitis, obesity, haemophilia, ulcers, anaemia, cancer, trauma, regeneration, and viral, bacterial or parasitic infections. The present sequence represents a NOVX protein

Sequence 1489 AA;

Query Match 12.4%; Score 217; DB 6; Length 1489;
 Best Local Similarity 28.2%; Pred. No. 3e-12;
 Matches 69; Conservative 46; Mismatches 112; Indels 18; Gaps 6;
 QY 75 NFPYSLEHLQTSYCGLVVRVDMRLCLSLRKLDSLHNHKKLPATIGDLHLQELNLDN 134
 Db 43 NFERTLEELYLDANQIEELPKQLENQALRKLSTPDNDLSNLTPTTASLVNKLKELDISKN 102
 QY 135 HLESF--SVALCHSTLOKSLRSLDLKKNKIKALPVQFCOLQELKLNKLDNELLQFPCKI 192
 Db 103 GVQEPFENIKCC-----KCLTITIEASVNPISKLPGFTQLNLTLQVLYNDAFLFELPANF 157
 QY 193 GOLINRLFLSAARNKLPFLPSEFNLSLEYLDLFGNTF-EQPKVLPVTKLQAPLTLLSS 251
 Db 158 GRUVKRLILELRENHLKTLPMWKLQAQLERLDLGNNEFSELPVLDQLQNLREL-WMDNN 216

QY 252 ARTLHNRPYGVSHIIPHLQD-LDTAKICVCGRFCLNSFIQGTTHNLSVAHTVVLV 310
 Db 217 ALQVLPDSIGKRLMLVYDMSKNRIETVDMISGCEALDLSNNMLQ-----QLP 268
 QY 311 DNLGG 315
 Db 269 DSIQG 273

RESULT 12
 ABB61551
 ID ABB61551 standard; protein; 431 AA.
 XX ABB61551;
 AC ABB61551;
 XX 26-MAR-2002 (first entry)
 DT Drosophila melanogaster polypeptide SEQ ID NO 11445.
 DE Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 KW Drosophila melanogaster.
 XX WO200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US009231.
 XX 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR N-PSDB; ABL05654.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX Disclosure; SEQ ID NO 11445; 21pp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
 CC sequences (ABU01840-ABU16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 431 AA;
 SQ Query Match 12.1%; Score 211.5; DB 4; Length 431;
 Best Local Similarity 31.6%; Pred. No. 1.8e-12;
 Matches 79; Conservative 39; Mismatches 105; Indels 27; Gaps 11;

QY 12 LQULFQA--ISSSLKGFISAMPLAHRGCVDFPVSTLTPVKTSFENFKTKMVI-TSKK 68
 Db 39 JEEFLFANHRDLPKNFFRLHRLKLGSL-DNEIGRLPP-----DIQNFENVELDVSRN 93
 QY 69 DYPLSNFFYSLHLOT-----SYCGLVRVDMRLCLSLKSLKLDLSHNHIKLPATIGD 122
 Db 94 DIP---DIPDDXHLQSLQVADESSNPIKLPSPGFSQKRLMTVLIGLNDMSLTLPADFGS 150
 QY 123 LIHQLQELNNDNEHSFVALCHSTLQKSLRSLSLDSKNNIKALPVQFCQLQELKNLKLDD 182

Db 151 LTQLESLELRNLAKHLPETISQLT---KLKELDLDGNEIEDLPYLGYPGLHLELWDH 207
 QY 183 NELIQFPCKIGQLNLNRLSAAARNKLPPLPSEFRL-SLEYLDLFGNTFEQKVLPP--VI 239
 Db 208 NQQLPPELGLLTKLYTLDVSENRLSELPNEISGLVSTLDDLAQNLLI---ALPDGIA 264
 QY 240 KQAPLTLLE 249
 Db 265 KL-SRLTILK 273

RESULT 13
 ABB63454
 ID ABB63454 standard; protein; 378 AA.
 XX ABB63454;
 AC ABB63454;
 XX 26-MAR-2002 (first entry)
 DT Drosophila melanogaster polypeptide SEQ ID NO 23154.
 DE Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 KW Drosophila melanogaster.
 XX WO200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US009231.
 XX 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR N-PSDB; ABL09557.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX Disclosure; SEQ ID NO 23154; 21pp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
 CC sequences (ABU01840-ABU16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 378 AA;
 SQ Query Match 11.9%; Score 207.5; DB 4; Length 378;
 Best Local Similarity 32.0%; Pred. No. 3.8e-12;
 Matches 66; Conservative 39; Mismatches 72; Indels 29; Gaps 8;

QY 82 HLQTSYCVGLVRVDMRLCLSLKSLKLDLSHNHIKLPATIGLHQLQELNNDNEHSFV 141
 Db 53 HLAGNNLSLPEVDIYV--LENLEFLDVSNNEIKELPPTGLLNLQQLNYSQNLTELPV 110
 QY 142 ALCHSTLQKSLRSLSKNNIKALPVQFCQLQELKNLKLDDNE-LIQFCKIGQLNIRF 200
 Db 111 EL--SGL--RNLEHNLGKQCFRFPVQLSECVRLNLSNDSNEALVHIPERISNLPMLQS 167

Qy 252 ARTILRN 258
Db : | |
356 DQDLRN 362

Search completed: June 15, 2004, 10:01:19
Job time : 62 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2004, 22:58:17 ; Search time 2398 Seconds
(without alignments)
15130.338 Million cell updates/sec

Title: US-10-009-557-34

Perfect score: 1215

Sequence: 1 gaagaactagcatgtatgtatgta.....ttagtttttaaaaaaaaaa 1215

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.*

1: em_estba.*

2: em_estum.*

3: em_estin.*

4: em_estu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_estc.*

9: gb_estl.*

10: gb_est2.*

11: gb_hic.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_man.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_pkp.*

27: em_gss_vrl.*

28: gb_gssi.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1091.2	89.8	1437	11	CNSLTI186
2	1050	86.4	1143	9	AL582092
3	956.6	78.7	989	9	AL582114
4	925.8	76.2	1155	9	AL560571

C	5	902.4	74.3	1133	13	BX405954
C	6	764.8	62.9	770	13	BX401074
	7	760.4	62.6	1150	13	BX333340
	8	759.2	62.5	948	13	BX370934
	9	758	62.4	791	13	BX370934
	10	733.6	60.4	901	13	BX370935
C	11	710	58.4	739	13	BX370935
C	12	705.2	58.0	1049	13	BX370935
C	13	703.8	57.9	722	13	BX370935
	14	700	57.6	891	13	BX370935
C	15	695.8	57.3	1165	13	BX370935
C	16	684	56.3	947	13	BX370935
C	17	660.4	54.4	972	13	BX370935
	18	649.6	53.5	1201	13	BX370935
	19	646.8	53.2	991	9	AL560705
	20	627.6	51.7	1032	12	SM554257
	21	610.4	50.2	1157	13	BX405955
	22	582.8	48.0	1072	12	BM476830
	23	538	44.3	613	9	AV717878
C	24	501.2	41.3	685	29	CE788486
	25	479.2	39.4	658	14	CF950469
C	26	476.6	39.2	572	9	AI739536
	27	462.6	38.1	1050	13	BX346615
C	28	458	37.7	483	12	BG940337
	29	449.8	37.0	874	14	CF265122
C	30	443.8	36.5	467	9	AA424400
C	31	433.8	35.7	466	9	AA812029
C	32	430.8	35.5	446	9	AA393227
C	33	430.8	35.5	645	14	CD365392
C	34	430	35.4	439	13	BU753893
C	35	430	35.4	561	14	CA421227
C	36	428.4	35.3	473	9	AA742244
	37	422.4	34.8	775	13	BQ053807
C	38	420.8	34.6	540	10	AM971618
C	39	419	34.5	656	13	BU620172
	40	412.2	33.9	801	12	B1258030
C	41	411.8	33.9	729	10	BG067404
	42	411.4	33.9	999	13	BU526707
	43	411	33.8	411	9	AA424474
C	44	409	33.7	512	9	AA648933
	45	406.2	33.4	1002	12	BM925426

ALIGNMENTS

RESULT 1

CNSLTI186

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CNSLTI186 1437 bp mRNA linear HTC 18-JUN-2003
human full-length cDNA clone CSODK007YE23 of HeLa cells of Homo sapiens (human).

BX248298

BX248298.1 GI:28207942

HTC.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1437)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ invitrogen Corporation 1600

Faraday Avenue

2 (bases 1 to 1437)

Genoscope.

Direct Submission

Submitted (04-FEB-2003) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

942 GATATTTATCCCTTGAATACTTGGATCTTTTGGAAATACTTTTGAACAACCAAGTCC 1001

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

```

Query Match      86.4%; Score 1050; DB 9; Length 1143;
Best Local Similarity 97.3%; Pred. No. 4.6e-231;
Matches 1084; Conservative 13; Mismatches 13; Indels 4; Gaps 3;

QY 59 TATTCAGGCCATTTCCAGCAGTTTAAAGGTTTCCTTCAGCTATCAGACTCGCTCATAG 118
DB 1110 TAAGDAAGCCATTTCCAGCAGTTTAAAGGTTT-CCTTTCATATGAACTGGCTCATAG 1053

QY 119 AGGCTGTATGTTGTATACACCACTTTCAACGCTCACACAGTGAAGACTTCAGAATTGA 178
DB 1052 AGGCTGTAAATGKRTACACCACTTTCAACGCTCACACAGTGAAGACTTCAGAATTGA 993

QY 179 AAACCTTTAAACTAAATGGTTTATCAGATCCAAAAGACTATCCTTAAGTAAGATTT 238
DB 992 AAACCTTTAAACTAAATGGTTTATCAGATCCAAAAGACTATCCTTAAGTAAGATTT 933

QY 239 TCCATATTCCTTGGAACTTTTCAAGCTTTTACTGTGGCTTGCAGTTGATATGG 298
DB 932 CCGATATTCCTTGGAACTTTTCAAGCTTTTACTGTGGCTTGCAGTTGATATGG 873

QY 299 TATGCTTTTGTAAAGCCCTTAGGAATTTAGACTTTGAGTCAACACCATATATAAAGCT 358
DB 872 TATGCTTTTGTAAAGCCCTTAGGAATTTAGACTTTGAGTCAACACCATATATAAAGCT 813

QY 359 TCCAGCTACATTTGGAGACCTTATACCTTCAAGAACTTAACTTCAATGATATCACT 418
DB 812 TCCAGCTACATTTGGAGACCTTATACCTTCAAGAACTTAACTTCAATGATATCACT 753

QY 419 GGAGTCAATTTAGTGTAGCTTTGTGTCATTTTACACTCCAGAGTCACTTCGGAGTTTGA 478
DB 752 GGAGTCAATTTAGTGTAGCTTTGTGTCATTTTACACTCCAGAGTCACTTCGGAGTTTGA 693

QY 479 CCTCAGCAGAACAAAATCAAGGCACTCCCTGTGCGAGTTTGGCCAGCTCCAGGAACCTAA 538
DB 633 CCTCAGCAGAACAAAATCAAGGCACTCCCTGTGCGAGTTTGGCCAGCTCCAGGAACCTAA 633

QY 539 GAATTTAAACTTCCAGATATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 598
DB 632 GAATTTAAACTTCCAGATATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 573

QY 599 AAACCTTGGCTTTTGTGAGCAGCTCGAAATAAGCTTCCATTTTTCCTAGTGAATTTAG 658
DB 572 AAACCTTGGCTTTTGTGAGCAGCTCGAAATAAGCTTCCATTTTTCCTAGTGAATTTAG 513

QY 659 AAATTTATCCCTTGAATCTTGGATCTTTTGGAAATCTTTTGGAAATCTTTTGGAAAT 718
DB 512 AAATTTATCCCTTGAATCTTGGATCTTTTGGAAATCTTTTGGAAATCTTTTGGAAAT 453

QY 719 TCCAGTAAATAAGCTGCAAGCACCATTAACTTTATTTGAATCTTTCTGACCAACCAATTT 778
DB 452 TCCAGTAAATAAGCTGCAAGCACCATTAACTTTATTTGAATCTTTCTGACCAACCAATTT 393

QY 779 ACATAATAGATTCATATAGGCTCTCATATCAATTCATTCATTCCTGCCAAGATTTTGA 838
DB 392 ACATAATAGATTCATATAGGCTCTCATATCAATTCATTCATTCCTGCCAAGATTTTGA 333

QY 839 TACCGCAAAATTTGTTTGTGGAGATTTCTGTGAACTTTCTGTGAACTTTCTGTGAACT 898
DB 332 TACCGCAAAATTTGTTTGTGGAGATTTCTGTGAACTTTCTGTGAACTTTCTGTGAACT 273

QY 899 TACCATGAATCTGATTTCTGTTTGGCCACACTGTGTTCTTAGTATATATTTGGTGTAC 958
DB 272 TACCATGAATCTGATTTCTGTTTGGCCACACTGTGTTCTTAGTATATATTTGGTGTAC 214

QY 959 TGAAGCACCTATTATCTTTCTTTATTTCTGTTCTTAGGCTGTTTATGTTAATTCCTCTGAT 1018
DB 213 TGAAGCACCTATTATCTTTCTTTATTTCTGTTCTTAGGCTGTTTATGTTAATTCCTCTGAT 154

```

```

QY 1019 GTTAAAGTAATGGTGGAGACCAAGAAAAGAAATTTCAATAACAGATCAGTTTGGGTGCA 1078
DB 153 GTTAAAGTAATGGTGGAGACCAAGAAAAGAAATTTCAATAACAGATCAGTTTGGGTGCA 94

QY 1079 TGTATGATTTTCCAGCGTCAAAATTTGGAGTAAGGAAGATTTCTGTATATCTTGTCTGAGAG 1138
DB 93 TGTATGATTTTCCAGCGTCAAAATTTGGAGTAAGGAAGATTTCTGTATATCTTGTCTGAGAG 34

QY 1139 GAGGATGATGATAGTACTTACTCATTTAGATGATCTC 1172
DB 33 GAGGATGATGATAGTACTTACTCATTTAGATGATCTC 1

```

```

RESULT 3
AL582114/c
LOCUS      AL582114.2 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DL004YC22 3-PRIME, mRNA sequence.
ACCESSION AL582114
VERSION    AL582114.2 GI:31320341
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 999)
AUTHORS     Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     On Feb 16, 2001 this sequence version replaced gi:12949780.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 7624.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DL004Bb1NP1&cluster=7624.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DL004BB-INP1.
            Location/Qualifiers
                1..999
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DL004YC22"
                /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
                /cell_line="RAMOS CELL LINE"
                /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
                25-NORMALIZED"
                /note="1st strand cDNA was primed with a NotI-oligo(dT)
                primer. Five prime end enriched, double-strand cDNA was
                digested with Not I and cloned into the Not I and EcoR V
                sites of the pCMVSPORT 6 vector. Library was normalized."

```

FEATURES

```

source
Query Match      78.7%; Score 956.6; DB 9; Length 989;
Best Local Similarity 98.7%; Pred. No. 1.4e-203;
Matches 976; Conservative 8; Mismatches 3; Indels 2; Gaps 2;

QY 179 AAATTTAAATCTAAAATGGTTTATCAGATCCAAAAGACTATCCTTAAGTAAGATTT 238
DB 989 AAATTTAAATCTAAAATGGTTTATCAGATCCAAAAGACTATCCTTAAGTAAGATTT 931

QY 239 TCCATATTCCTTGGAACTTTTCAAGCTTTTACTGTGGCTTGTCCAGTTGATATGG 298
DB 930 TCCATATTCCTTGGAACTTTTCAAGCTTTTACTGTGGCTTGTCCAGTTGATATGG 871

QY 299 TATGCTTTGCTTAAAGCCCTTAGGAATTTAGCTTGGATCAGACCATAT-AAAAAGC 357

```

ORIGIN

```

Query Match      78.7%; Score 956.6; DB 9; Length 989;
Best Local Similarity 98.7%; Pred. No. 1.4e-203;
Matches 976; Conservative 8; Mismatches 3; Indels 2; Gaps 2;

QY 179 AAATTTAAATCTAAAATGGTTTATCAGATCCAAAAGACTATCCTTAAGTAAGATTT 238
DB 989 AAATTTAAATCTAAAATGGTTTATCAGATCCAAAAGACTATCCTTAAGTAAGATTT 931

QY 239 TCCATATTCCTTGGAACTTTTCAAGCTTTTACTGTGGCTTGTCCAGTTGATATGG 298
DB 930 TCCATATTCCTTGGAACTTTTCAAGCTTTTACTGTGGCTTGTCCAGTTGATATGG 871

QY 299 TATGCTTTGCTTAAAGCCCTTAGGAATTTAGCTTGGATCAGACCATAT-AAAAAGC 357

```

```
Db      870  TATGCTTTGCTTAAAGAGCCTTAGGAAATTAGACTTGAGTCACACCACTATATAAAAAAGC 811
Qy      358  TTCAGCTACAAATTGGAGACCTCTACACCTTCAAGAACTTAACTGATGACATCACT 417
Db      810  TTCAGCTACAAATTGGAGACCTCTACACCTTCAAGAACTTAACTGATGACATCACT 751
Qy      418  TGGAGTCATTTAGTGTAGCTTGTCTATCTACACTCCAGAACTCACTTCGGAGTTGG 477
Db      750  TGGAGTCATTTAGTGTAGCTTGTCTATCTACACTCCAGAACTCACTTCGGAGTTGG 691
Qy      478  ACCTCAGCAAGAACAAATCAAGGACCTCCCTGTGCGAGTTTCCAGCTCCAGGAACCTA 537
Db      690  ACCTCAGCAAGAACAAATCAAGGACCTCCCTGTGCGAGTTTCCAGCTCCAGGAACCTA 631
Qy      538  AGAATTTAAATCTGACGATATGAATGATTCAATTTCCCTTGAAGATAGACAACTAA 597
Db      630  AGAATTTAAATCTGACGATATGAATGATTCAATTTCCCTTGAAGATAGACAACTAA 571
Qy      598  TAAACCTTCGCTTTTGTGACGAGCTCGAAATAAGCTTCCATTTTTCCTAGTGAATTA 657
Db      570  TAAACCTTCGCTTTTGTGACGAGCTCGAAATAAGCTTCCATTTTTCCTAGTGAATTA 511
Qy      658  GAAATTTATCCCTTGATACCTTGTGATCTTTTGGAAATACCTTTGAAACCAACCAAGTCC 717
Db      510  GAAATTTATCCCTTGATACCTTGTGATCTTTTGGAAATACCTTTGAAACCAACCAAGTCC 451
Qy      718  TTCAGTAATAAGCTGCAAGCACCATTAACTTTATTCGAAATCTTCTGCAAGCACTAT 777
Db      450  TTCAGTAATAAGCTGCAAGCACCATTAACTTTATTCGAAATCTTCTGCAAGCACTAT 391
Qy      778  TACATAATAGGAATCCATATGCTCTCATATCATTCATTCATTCATCTGCGCAAGATTGG 837
Db      390  TACATAATAGGAATCCATATGCTCTCATATCATTCATTCATTCATCTGCGCAAGATTGG 331
Qy      838  ATACCGCAAAATTTGTTTGTGGAGAACTCTGTCGAACTCTTCATTCATTCATTCATTC 897
Db      330  ATACCGCAAAATTTGTTTGTGGAGAACTCTGTCGAACTCTTCATTCATTCATTCATTC 271
Qy      898  CTACATGAATCTGCAATCTCTGTCGCCACACTGTGGTCTTAGTAGATAAATTTGGGTGTA 957
Db      270  CTACATGAATCTGCAATCTCTGTCGCCACACTGTGGTCTTAGTAGATAAATTTGGGTGTA 211
Qy      958  CTGAAGCACTATTAATCTCTTATTTCTCTTCTAGGCTGTTAATCTCTGATA 1017
Db      210  CTGAAGCACTATTAATCTCTTATTTCTCTTCTAGGCTGTTAATCTCTGATA 151
Qy      1018  TGTAAAGTAATGGTGAGACCAAGAAAGAAATTTCAATACAGATCAGTTTGGGTGC 1077
Db      150  TGTAAAGTAATGGTGAGACCAAGAAAGAAATTTCAATACAGATCAGTTTGGGTGC 91
Qy      1078  ATGTATGATTTTGCAGGCTCAAAATGGAGTAGGAAGAGATTTCTGTATACCTTGTCTGGAGA 1137
Db      90  ATGTATGATTTTGCAGGCTCAAAATGGAGTAGGAAGAGATTTCTGTATACCTTGTCTGGAGA 31
Qy      1138  GGAGGAATGTGTATAGTTACTCAATTTAGA 1166
Db      30  GGAGGAATGTGTATAGTTACTCAATTTAGA 2

RESULT 4
LOCUS   AL560571
DEFINITION Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
Homo sapiens cDNA clone CS0DL003VP23 5-PRIME, mRNA sequence.
ACCESSION AL560571
VERSION   AL560571.2 GI:31284701
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1155)
AUTHORS  Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
```

```
TITLE
JOURNAL
COMMENT
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12907161.
Contact: Centre National de Sequencage
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7624.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DL003CC12QP1&cluster=7624.f. Contact :
Feng liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DL003CC12QP1.
FEATURES
Location/Qualifiers
1..1155
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL003VP23"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

```
Query Match      76.2%; Score 925.8; DB 9; Length 1155;
Best Local Similarity 96.0%; Pred. No. 1.7e-202;
Matches 982; Conservative 17; Mismatches 19; Indels 5; Gaps 5;

Qy      58  TTATTCAGCGCATTTCCAGCAGTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCAT 117
Db      124  TAAGTAAGGCCATTTCCAGCAGTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCAT 183
Qy      118  GAGGCTGTAATGTTGATACACCGCTTCAACGCTCACACCACTGAAGACTTCAGAAATTG 177
Db      184  GAGGCTGTAATGTTGATACACCGCTTCAACGCTCACACCACTGAAGACTTCAGAAATTG 243
Qy      178  AAAAATTTAAAGCTAAATGTTATCACATCAAAAAGAGACTATCCTCTAAGTAAGAATT 237
Db      244  AAAAATTTAAAGCTAAATGTTATCACATCAAAAAGAGACTATCCTCTAAGTAAGAATT 303
Qy      238  TTCCATATTCCTTGGACATCTTCAGACTTCTTACTGTGGGCTTGTCCGAGTTGATGC 297
Db      304  TTCCATATTCCTTGGACATCTTCAGACTTCTTACTGTGGGCTTGTCCGAGTTGATGC 363
Qy      298  GTATGCTTTGCTTAAAGAGCCTTAGGAAATTAGACTTCAGTCACAACTATATAAAAGC 357
Db      364  GTATGCTTTGCTTAAAGAGCCTTAGGAAATTAGACTTCAGTCACAACTATATAAAAGC 423
Qy      358  TTCCAGCTACAAATTTGGAGACCTCTATACACTTCAAGAACTTAACTGATGACATCACT 417
Db      424  TTCCAGCTACAAATTTGGAGACCTCTATACACTTCAAGAACTTAACTGATGACATCACT 483
Qy      418  TGGAGTCATTTAGTGTAGCTTGTCTATCTACACTCCAGAACTCACTTCGGAGTTGG 477
Db      484  TGGAGTCATTTAGTGTAGCTTGTCTATCTACACTCCAGAACTCACTTCGGAGTTGG 543
Qy      478  ACCTCAGCAAGAACAAATCAAGGACCTCCCTGTGCGAGTTTCCAGCTCCAGGAACCTA 537
Db      544  ACCTCAGCAAGAACAAATCAAGGACCTCCCTGTGCGAGTTTCCAGCTCCAGGAACCTA 603
Qy      538  AGAATTTAAATCTGACGATATGAATGATTCAATTTCCCTTGAAGATAGACAACTAA 597
Db      604  AGAATTTAAATCTGACGATATGAATGATTCAATTTCCCTTGAAGATAGACAACTAA 663
Qy      598  TAAACCTTCGCTTTTGTGACGAGCTCGAAATAAGCTTCCATTTTTCCTAGTGAATTA 657
```


QY 1139 GAGGAATGT 1147
DB 17 GAGGAATKT 9

RESULT 6
LOCUS BX401074/c
DEFINITION BX401074 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
ACCESSION BX401074.1 GI:30614440
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7624.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK007AC12NP1&cluster=7624.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DK007AC12NP1.
Location/Qualifiers
1. .770
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK007YB23"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source
Query Match 62.9%; Score 764.8; DB 13; Length 770;
Best Local Similarity 99.4%; Pred. No. 1.9e-165;
Matches 765; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 384 CACCTTCAAGAACCTTAACCTGAAATGACATCACTTGGAGTCACTTAGTGTACCTTGTGT 443
DB 770 CACCTTCAAGAACCTTAACCTGAAATGACATCACTTGGAGTCACTTAGTGTACCTTGTGT 711

QY 444 CATCTACACTCCAGAACTCACTTCGGAGTTGGACCTCAGCAAGCAAAATCAAGCA 503
DB 710 CATCTACACTCCAGAACTCACTTCGGAGTTGGACCTCAGCAAGCAAAATCAAGCA 651

QY 504 CTCCTGTGCGAGTTTGGCAGCTCCAGAACTTAAGAAATTAATAAATTTGACGATATGAA 563
DB 650 CTCCTGTGCGAGTTTGGCAGCTCCAGAACTTAAGAAATTTAAATCTTGACGATATGAA 591

QY 564 TTGATTCGAATTCCTTGCAGATAGGACACTTAATACCTTCGCTTTTGTGACAGCT 623
DB 590 TTGATTCGAATTCCTTGCAGATAGGACACTTAATACCTTCGCTTTTGTGACAGCT 531

QY 624 CGAATAAGCTTCCATTTTTCGCTAGTGAATTTAGAAATTTATCCCTTGAATCTTGGAT 683
DB 530 CGAATAAGCTTCCATTTTTCGCTAGTGAATTTAGAAATTTATCCCTTGAATCTTGGAT 471

QY 684 CTTTGTGGAATACTTTTGAACCAACCAAAAGTCTCTCCAGTAAATAGAGCTGCAGACCA 743
DB 470 CTTTGTGGAATACTTTTGAACCAACCAAAAGTCTCTCCAGTAAATAGAGCTGCAGACCA 411

QY 744 TTAACCTTTATTTGGAATCTTCTGCACGAACCATATTACATAATAGGATTCATATGGCTCT 803
DB 410 TTAACCTTTATTTGGAATCTTCTGCACGAACCATATTACATAATAGGATTCATATGGCTCT 351

QY 804 CATATCATTTCCATTCATCTCTGCCAAGATTGGATACCGCAAAAATTTGTTGTGGA 863
DB 350 CATATCATTTCCATTCATCTCTGCCAAGATTGGATACCGCAAAAATTTGTTGTGGA 291

QY 864 AGATTTCTGTCTGAATCTCTTTCATTCAAGGAACCTACTACCAATGAATCTGATTTGCC 923
DB 290 AGATTTCTGTCTGAATCTCTTTCATTCAAGGAACCTACTACCAATGAATCTGATTTGCC 231

QY 924 CACACTGTGGTCTTCTAGTAGATAATTTGGGTGGTACTCTGAAGCACTTATCTCTTATTTTC 983
DB 230 CACACTGTGGTCTTCTAGTAGATAATTTGGGTGGTACTCTGAAGCACTTATCTCTTATTTTC 171

QY 984 TGTCTCTAGGCTGTATGTAAATTCCTCTGATATCTTAAAGTAATGGGTGAGACAGAA 1043
DB 170 TGTCTCTAGGCTGTATGTAAATTCCTCTGATATCTTAAAGTAATGGGTGAGACAGAA 111

QY 1044 AAAGAAATTTCAATAACACATCAGTTTGGGTGCATGTATGATTTTGCAGCGTCAATTTG 1103
DB 110 AAAGAAATTTCAATAACACATCAGTTTGGGTGCATGTATGATTTTGCAGCGTCAATTTG 51

QY 1104 GAGTAAGGAAGATTTCTGTATATCTTCTGAGAGGAGGAGGAATGTGTATAG 1153
DB 50 GAGTAAGGAAGATTTCTGTATATCTTCTGAGAGGAGGAGGAATGTGTATAG 1

RESULT 7
LOCUS BX333340
DEFINITION BX333340 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
ACCESSION BX333340
VERSION BX333340.1 GI:30308198
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7624.f For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DK024AB03QP1
cluster=7624.f. Contact : Feng Liang Email : fliang@lifetech.com
URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DK024AB03QP1.
Location/Qualifiers
1. .115C
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK024YC05"
/tissue type="NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source
1. .115C

ORIGIN

```

Query Match      62.6%; Score 760.4; DB 13; Length 1150;
Best Local Similarity 99.9%; Pred. No. 1.9e-164;
Matches 761; Conservative 0; Mismatches 1; Indels 3; Gaps 0;

QY 454 TCAGAGTACATTCGGAGTTGGAGCTCAGCAAGAACAAATCAAGGCACTCCCTGTGC 513
DB 122 AGTTTGGCCAGCTCCAGGAACTTAAGATTTAAACTTGAAGTAATGAATCAAT 181
QY 574 TTCCCTGCAAGTAGACAACTTAATAAACTTGGTTCAGCAGCTCGAAATAAGC 633
DB 182 TTCCCTGCAAGTAGACAACTTAATAAACTTGGTTCAGCAGCTCGAAATAAGC 241
QY 634 TTCCCTGCAAGTAGACAACTTAATAAACTTGGTTCAGCAGCTCGAAATAAGC 693
DB 242 TTCCCTGCAAGTAGACAACTTAATAAACTTGGTTCAGCAGCTCGAAATAAGC 301
QY 694 ATACTTTTGAACAACCAAAAGCTCTTCCAGTAATAAGCTGCAAGCACTTAATCAAT 753
DB 302 ATACTTTTGAACAACCAAAAGCTCTTCCAGTAATAAGCTGCAAGCACTTAATCAAT 361
QY 754 TGAATCTTTCAGCAAGCACTTAATCAATAAGTAATCCATATGCTCTCATATCAATTC 813
DB 362 TGAATCTTTCAGCAAGCACTTAATCAATAAGTAATCCATATGCTCTCATATCAATTC 421
QY 814 CATTCATCTTCGCAAGATTTGGATACCGCAAAATTTGCTTGTGGAAGATCTGTGC 873
DB 422 CATTCATCTTCGCAAGATTTGGATACCGCAAAATTTGCTTGTGGAAGATCTGTGC 481
QY 874 TGAATCTTTCAGCAAGCACTTAATCAATAAGTAATCCATATGCTCTCATATCAATTC 933
DB 482 TGAATCTTTCAGCAAGCACTTAATCAATAAGTAATCCATATGCTCTCATATCAATTC 541
QY 934 TCTTAGTAGAATAATTTGGTGTGCTAGAGCACTTAATCTTCTTATTTCTGTCTAG 993
DB 542 TCTTAGTAGAATAATTTGGTGTGCTAGAGCACTTAATCTTCTTATTTCTGTCTAG 601
QY 994 GCTGTATGTAAATTCCTCTGATATGTTAAAGTAATGGTGTGAGACCAAGAAAGATTT 1053
DB 602 GCTGTATGTAAATTCCTCTGATATGTTAAAGTAATGGTGTGAGACCAAGAAAGATTT 661
QY 1054 CAATAACAGATCAGTTGGGTGCTATGATGATTTTGCAGCGCTCAATTTGGAGTAAAGGA 1113
DB 662 CAATAACAGATCAGTTGGGTGCTATGATGATTTTGCAGCGCTCAATTTGGAGTAAAGGA 721
QY 1114 AGATTTCTGTATCTGCTGAGAGAGCAATGCTATAGTACTCATTTAGATCACTCC 1173
DB 722 AGATTTCTGTATCTGCTGAGAGAGCAATGCTATAGTACTCATTTAGATCACTCC 781
QY 1174 AAAAATTTTATTAACCAATTTTGTGTTTAAAAAATTTTAAAAAATTTTAAAAA 1215
DB 782 AAAAATTTTATTAACCAATTTTGTGTTTAAAAAATTTTAAAAAATTTTAAAAA 823

BX370934      948 bp  mRNA  linear  EST 08-MAY-2003
BX370934 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS0024Y05 5-PRIME, mRNA sequence.
ACCESSION     BX370934.1 GI:30457965
VERSION       BX370934
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 948)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization

```

```

JOURNAL COMMENT
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 131 91006 EVRY cedex - France
Email: seq@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7624.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAH004Z090P1&cluster=7624.f. Contact :
Peng Liang Email : filiang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1500
Faraday Avenue Genoscope sequence ID : CS0BAH004Z090P1.
FEATURES
source
1..948
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0024Y05"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match      62.5%; Score 759.2; DB 13; Length 948;
Best Local Similarity 99.8%; Pred. No. 3.6e-164;
Matches 759; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 454 TCAGAGTACATTCGGAGTTGGAGCTCAGCAAGAACAAATCAAGGCACTCCCTGTGC 513
DB 66 TCAGAGTACATTCGGAGTTGGAGCTCAGCAAGAACAAATCAAGGCACTCCCTGTGC 125
QY 514 AGTTTGGCCAGCTCCAGGAACTTAAGANTTTAAACTTGAAGTAATGAATCAAT 573
DB 126 AGTTTGGCCAGCTCCAGGAACTTAAGANTTTAAACTTGAAGTAATGAATCAAT 185
QY 574 TTCCCTGCAAGTAGACAACTTAATAAACTTGGTTCAGCAGCTCGAAATAAGC 633
DB 186 TTCCCTGCAAGTAGACAACTTAATAAACTTGGTTCAGCAGCTCGAAATAAGC 245
QY 634 TTCCCTGCAAGTAGACAACTTAATAAACTTGGTTCAGCAGCTCGAAATAAGC 693
DB 246 TTCCCTGCAAGTAGACAACTTAATAAACTTGGTTCAGCAGCTCGAAATAAGC 305
QY 694 ATACTTTTGAACAACCAAAAGCTCTTCCAGTAATAAGCTGCAAGCACTTAATCAAT 753
DB 306 ATACTTTTGAACAACCAAAAGCTCTTCCAGTAATAAGCTGCAAGCACTTAATCAAT 365
QY 754 TGAATCTTTCAGCAAGCACTTAATCAATAAGTAATCCATATGCTCTCATATCAATTC 813
DB 366 TGAATCTTTCAGCAAGCACTTAATCAATAAGTAATCCATATGCTCTCATATCAATTC 425
QY 814 CATTCATCTTCGCAAGATTTGGATACCGCAAAATTTGCTTGTGGAAGATCTGTGC 873
DB 426 CATTCATCTTCGCAAGATTTGGATACCGCAAAATTTGCTTGTGGAAGATCTGTGC 485
QY 874 TGAATCTTTCAGCAAGCACTTAATCAATAAGTAATCCATATGCTCTCATATCAATTC 933
DB 486 TGAATCTTTCAGCAAGCACTTAATCAATAAGTAATCCATATGCTCTCATATCAATTC 545
QY 934 TCTTAGTAGAATAATTTGGTGTGCTAGAGCACTTAATCTTCTTATTTCTGTCTAG 993
DB 546 TCTTAGTAGAATAATTTGGTGTGCTAGAGCACTTAATCTTCTTATTTCTGTCTAG 605
QY 994 GCTGTATGTAAATTCCTCTGATATGTTAAAGTAATGGTGTGAGACCAAGAAAGATTT 1053
DB 606 GCTGTATGTAAATTCCTCTGATATGTTAAAGTAATGGTGTGAGACCAAGAAAGATTT 665
QY 1054 CAATAACAGATCAGTTGGGTGCTATGATGATTTTGCAGCGCTCAATTTGGAGTAAAGGA 1113

```

Db 666 CAATAACAGATCAGTTGGGGTGCATGATGATTTTGCAGCGTCAAAATTGGAGTDAAGGA 725

QY 1114 AGATTCTGTATACCTCTCGAGAGGAGGATGATAGTTACTTCATTTAGATGACTCC 1173

Db 726 AGATTCTGTATACCTCTCGAGAGGAGGATGATAGTTACTTCATTTAGATGACTCC 785

QY 1174 AAAACCTTTTATTAACCAATTTTGTAGTTTAAAAAAGAAAA 1215

Db 786 AAAACCTTTTATTAACCAATTTTGTAGTTTAAAAAAGAAAA 827

RESULT 9
 BU852149
 LOCUS
 DEFINITION AGENCOURT 10402016 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6617621
 5', mRNA sequence.

ACCESSION BU852149.1 GI:24037112
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 JOURNAL 1 (bases 1 to 791)
 COMMENT NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgraphs-remail.nih.gov
 Tissue Procurement: CLONTECH
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LICM2863 row: d column: 05
 High quality sequence stop: 588.

FEATURES
 Location/Qualifiers
 1..791
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6617621"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH MGC 82"
 /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1:
 SfiI (ggccattatggcc); Site 2: SfiI (ggccattatggcc); 5' and
 3' adaptors were used in cloning as follows: 5' adaptor
 sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
 5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size
 1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
 inserts by PCR. This library was enriched for full-length
 clones and was constructed by Clontech Laboratories (Palo
 Alto, CA)."

ORIGIN
 Query Match 62.4%; Score 758; DB 13; Length 791;
 Best Local Similarity 99.2%; Pred. No. 6.9e-164;
 Matches 761; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 449 TCACCTCCAGAGTCACTTCGGAGTTTGGACCTCAGCAAGACAAATCAAGGCACTCCC 508

Db 3 TCACCTCCAGAGTCACTTCGGAGTTTGGACCTCAGCAAGACAAATCAAGGCACTCCC 62

QY 509 TGTGCGAGTTTCCAGCTCCAGAACTTAAAGATTTAAACCTTGACGATTAATGAT 568

Db 63 TGTGCGAGTTTCCAGCTCCAGAACTTAAAGATTTAAACCTTGACGATTAATGAT 122

QY 569 TCAATTTCTTTCAGATAGGACAACTTAATAACCTTTCGCTTTTGTGACGACCTCGAA 628

Db 123 TCAATTTCTTTCAGATAGGACAACTTAATAACCTTTCGCTTTTGTGACGACCTCGAA 182

QY 629 TAAGCTCCATTTTGGCTAGTGAATTTAGAAATTTATCCCTGGAATACTTGGTCTTTT 688

Db 183 TAAGCTCCATTTTGGCTAGTGAATTTAGAAATTTATCCCTGGAATACTTGGTCTTTT 242

QY 689 TGGAAATACCTTTTGAACCAACCAAAAGTCTTCCAGTAATAAAGCTGCAAGCAACCAATTAAC 748

Db 243 TGGAAATACCTTTTGAACCAACCAAAAGTCTTCCAGTAATAAAGCTGCAAGCAACCAATTAAC 302

QY 749 TTTATTCGAATCTTCTGCAAGAACCATATTACATAATAGGATTCATATGCTCTCATAT 808

Db 303 TTTATTCGAATCTTCTGCAAGAACCATATTACATAATAGGATTCATATGCTCTCATAT 362

QY 809 CATTCACATTCATCTCTGCCAAGATTTGGATACCGCAAAATTTGTCTTGTGGAAGATT 868

Db 363 CATTCACATTCATCTCTGCCAAGATTTGGATACCGCAAAATTTGTCTTGTGGAAGATT 422

QY 869 CTGTCTGAACCTCTTTCATTCAGGAACCTACTACCATGAATCTGCATTCCTGTGTGCCACAC 928

Db 423 CTGTCTGAACCTCTTTCATTCAGGAACCTACTACCATGAATCTGCATTCCTGTGTGCCACAC 482

QY 929 TGTGCTCTTAGTAGTAATTTGGTGGTACTGAAGCACTTATATCTCTTATTTCTGTTTC 988

Db 483 TGTGCTCTTAGTAGTAATTTGGTGGTACTGAAGCACTTATATCTCTTATTTCTGTTTC 542

QY 989 TCTAGGCTGTTTATGCTTAATTTCTCTGATATGTTTAAAGTAATGGTGGAGCAAGAAAGA 1048

Db 543 TCTAGGCTGTTTATGCTTAATTTCTCTGATATGTTTAAAGTAATGGTGGAGCAAGAAAGA 602

QY 1049 AATTTCAATTAACAGATCAGTTTGGGGTGCATGATGATTAATTTGACGCTCAAAATGGAGTA 1108

Db 603 AATTTCAATTAACAGATCAGTTTGGGGTGCATGATGATTAATTTGACGCTCAAAATGGAGTA 662

QY 1109 AGGAGATTTCTGTATCTTGTGAGAGAGAGAAATCTGTATAGTTACTCATTTAGATG 1168

Db 663 AGGAGATTTCTGTATCTTGTGAGAGAGAGAAATCTGTATAGTTACTCATTTAGATG 722

QY 1169 ACTCCAAACCTTTTATTAACCAATTTTATTTTAAAAAAGAAAA 1215

Db 723 ACTCCAAACCTTTTATTAACCAATTTTATTTTAAAAAAGAAAA 769

RESULT 10
 BU852149
 LOCUS
 DEFINITION BX370935 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 cDNA clone CS0DC024YC05 5-PRIME, mRNA sequence.

ACCESSION BX370935
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 JOURNAL 1 (bases 1 to 901)
 COMMENT Full-length cDNA libraries and normalization
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 7624.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0BAH0342E09RM1&cluster=7624.f. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0BAH004ZE09RM1.

FEATURES
 Location/Qualifiers
 1..901
 /organism="Homo sapiens"

```

/ncbi_type="mRNA"
/db_xref="taxon:9606"
/clone="CS00C024YC05"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="First strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 60.4%; Score 733.6; DB 13; Length 901;
Best Local Similarity 97.6%; Pred. No. 2.8e-158;
Matches 742; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 454 TCCAGAGTCACCTCGAGTTGGACCTCAGCAGAACCAAAATCAAGGCACTCCCTGTGC 513
Db 76 TCCAGAGTCACCTTCGAGTTGGACCTCAGCAGAACCAAAATCAAGGCACTCCCTGTGC 135
Qy 514 AGTTTTCAGCTCCAGGCACTTGAAGATTTAAACCTTCAAGCTATGATTCAT 573
Db 136 AGTTTTCAGCTCCAGGCACTTGAAGATTTAAACCTTCAAGCTATGATTCAT 195
Qy 574 TTCCCTTGCAGATAGCACTAAATAACCTTCCTTTTGTTCAGCAGCTCGAAATAAGC 633
Db 196 TTCCCTTGCAGATAGCACTAAATAACCTTCCTTTTGTTCAGCAGCTCGAAATAAGC 255
Qy 634 TTCCATTTTGGCTAGTGAATTTAGAAATTTATCCCTTGAATCTTGGATTTGGAA 693
Db 256 TTCCATTTTGGCTAGTGAATTTAGAAATTTATCCCTTGAATCTTGGATTTGGAA 315
Qy 694 ATACATTTTGAACACCAAGTCCCTCCAGTAAATTAAGCTGCAAGCACCATTATCT 753
Db 316 ATACATTTTGAACACCAAGTCCCTCCAGTAAATTAAGCTGCAAGCACCATTATCT 375
Qy 754 TGGAACTTCTTCAGCAACCATATACATAATAGAGATTCATATGCTCTCATATCATTC 813
Db 376 TGGAACTTCTTCAGCAACCATATACATAATAGAGATTCATATGCTCTCATATCATTC 435
Qy 814 CATTCATCTTCGCAAGATTTGGATACCGCAAAATTTGTGTTTGGAGATTCCTGC 873
Db 436 CATTCATCTTCGCAAGATTTGGATACCGCAAAATTTGTGTTTGGAGATTCCTGC 495
Qy 874 TGAATCTTTCATCAAGAACTACTACCATGATCTGCATTCCTGTCGCCACACCTGG 933
Db 496 TGAATCTTTCATCAAGAACTACTACCATGATCTGCATTCCTGTCGCCACACCTGG 555
Qy 934 TCTTAGTAGAATATTTGGGTGTGTCAGAACCCATTAATCTTATTTCTGTTCTCAG 993
Db 556 TCTTAGTAGAATATTTGGGTGTGTCAGAACCCATTAATCTTATTTCTGTTCTCAG 615
Qy 994 GCTGTATCTTAATTCCTCTGATATGTTAAAGTAAATGGGTGAGACCAGAAAAGAACTT 1053
Db 616 GCTGTATCTTAATTCCTCTGATATGTTAAAGTAAATGGGTGAGACCAGAAAAGAACTT 675
Qy 1054 CAATAACAGATCAGTTTGGGTGTCATGATATTTTGCAGCGTCAATTTGGAGTAGGGA 1113
Db 676 GAATAACAGATCAGTTTGGGTGTCATGATATTTTGCAGCGTCAATTTGGAGTAGGGA 735
Qy 1114 AGATTTCTGTATCTTCTGAGAGGAGGAATGTGTATAGTTACTCATTTAGATGCTCC 1173
Db 736 AGATTTGTGTATCTTCTGAGAGGAGGAATGTGTATAGTTACTCATTTAGATGCTCG 795
Qy 1174 AAAACCTTTTATTAACCAATTTAGTTTAAAAAANA 1213
Db 796 AAAACCTTTTATTAACCAATTTTGGTTGTAAAAAANA 835

RESULT 11
BU621671/c
LOCUS
DEFINITION
UI-H-FL1-bga-b-24-0-UI.s1 NCI CGAP FL1 Homo sapiens cDNA clone
UI-H-FL1-bga-b-24-0-UI 3', mRNA sequence.

```

```

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
ORIGIN
Query Match 58.4%; Score 710; DB 13; Length 739;
Best Local Similarity 99.6%; Pred. No. 7.6e-153;
Matches 732; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Qy 481 TCAGCAAGAACAAAATCAAGGCACTCCCTGTGAGTTTGCAGCTCCAGGACTTAAGA 540
Db 739 TCAGCAAGAACAAAATCAA-GCACTCCCTGTGAG-TTTGCCAGCTNCAGGACTTAAGA 682
Qy 541 ATTAAAGCTTCAGATATGATTCATTTCAATTCCTTCCAGATAGGACACTAATA 600
Db 681 ATTAAAGCTTCAGATATGATTCATTTCAATTCCTTCCAGATAGGACACTAATA 652
Qy 601 ACCTTCGCTTTTGTCCAGCAGCTCGAAATAAGCTTCATTTTGCCTAGTGAATTAGAA 663
Db 621 ACCTTCGCTTTTGTCCAGCAGCTCGAAATAAGCTTCATTTTGCCTAGTGAATTAGAA 562
Qy 661 ATTTATCCCTGAATACTTGGACCTTTTGGAAATACTTTTGAAACCAACCAAGTCTTC 720
Db 661 ATTTATCCCTGAATACTTGGACCTTTTGGAAATACTTTTGAAACCAACCAAGTCTTC 720

/ncbi_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FL1-bga-b-24-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FL1"
/note="Organ: Chondrosarcoma; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FL1 is a normalized cDNA library derived from a pool of mRNA obtained from 4 cell lines from grade III chondrosarcoma tissues. The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GAGTCCGCTG. The cell lines were provided by Dr. James Martin from the University of Iowa.
TAG_ISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG_LIB=UI-H-FL1
TAG_SEQ=GAGTCCGCTG"

```

```

BU621671
BU621671.1 GI:23287886
EST.
Homo sapiens (human);
Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 739)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-43, >AT-rich#Low complexity (matched complement);
Seq primer: M13 FORWARD
POLYA=Yes.

```

```

Location/Qualifiers
1..739
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FL1-bga-b-24-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FL1"
/note="Organ: Chondrosarcoma; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FL1 is a normalized cDNA library derived from a pool of mRNA obtained from 4 cell lines from grade III chondrosarcoma tissues. The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GAGTCCGCTG. The cell lines were provided by Dr. James Martin from the University of Iowa.
TAG_ISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG_LIB=UI-H-FL1
TAG_SEQ=GAGTCCGCTG"

```

```

Db      561 ATTATCCCTTGAATCTTGGATCTTTTGGAAATACCTTTTGAACCAACCAAAAGTCCTTC 502
Qy      721 CAGTAATAAGAGCTGCAAGCACCATTATCTTATTTGGATCTTCTGCGAGCAACATATATAC 783
Db      501 CAGTAATAAGAGCTGCAAGCACCATTATCTTATTTGGATCTTCTGCGAGCAACATATATAC 442
Qy      781 ATAATAGGATTCATATAGGCTCTCATATCAATTCATTCATTCCTCTGCGAGCAACATATAT 840
Db      441 ATAATAGGATTCATATAGGCTCTCATATCAATTCATTCCTCTGCGAGCAACATATAT 382
Qy      841 CCGCAAAATTTGGTTTGGAGATCTCTGCTGAACCTCTTTCATTCATTCAGAACTACTA 900
Db      381 CCGCAAAATTTGGTTTGGAGATCTCTGCTGAACCTCTTTCATTCATTCAGAACTACTA 322
Qy      901 CCATGAATCTGCATCTGTTTGGCCACACTGCTGTAGTAGATAAATTTGGTGGTACTG 960
Db      321 CCATGAATCTGCATCTGTTTGGCCACACTGCTGTAGTAGATAAATTTGGTGGTACTG 262
Qy      961 AAGCACTATTAATCTCTTATTTCTGTTCTCTAGGCTGTATGTTAAATTCCTCTCATATGT 1020
Db      261 AAGCACTATTAATCTCTTATTTCTGTTCTCTAGGCTGTATGTTAAATTCCTCTCATATGT 202
Qy      1021 TAAAGTAATGGTGGAGACCAAGAAAGAAATTTCAATTAACAGATCAGTTTGGGTGGATG 1080
Db      201 TAAAGTAATGGTGGAGACCAAGAAAGAAATTTCAATTAACAGATCAGTTTGGGTGGATG 142
Qy      1081 TATGATTTTGCAGCTCAAAATTTGGAGTAAGGAGATTTCTGTATATCTGCTGGAGAGGA 1140
Db      141 TATGATTTTGCAGCTCAAAATTTGGAGTAAGGAGATTTCTGTATATCTGCTGGAGAGGA 82
Qy      1141 GGAATGTATATGTTACTCTATCTATTTAGATGATCTCCAAACTTTTATTAACCAATTTTGT 1200
Db      81 GGAATGTATATGTTACTCTATCTATTTAGATGATCTCCAAACTTTTATTAACCAATTTTGT 22
Qy      1201 TTTAAAAAABAAAA 1215
Db      21 TTTAAAAAABAAAA 7

```

```

RESULT 12
LOCUS   BX346614/c
DEFINITION BX346614 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
          cDNA clone CS0DB005YL20 3-PRIME, mRNA sequence.
ACCESSION BX346614
VERSION   BX346614.1 GI:30369111
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 1049)
          Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
          Full-length cDNA libraries and normalization
          Unpublished (2001)
          Contact: Genoscope
          Genoscope - Centre National de Sequencage
          BP 191 91006 EVRY cedex - France
          Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
          Library was constructed by Life Technologies, a division of
          Invitrogen. This sequence belongs to sequence cluster 7624.f For
          more information about this cluster, see
          http://www.genoscope.cns.fr/
          cgi-bin/cluster.cgi?seq=CS0AS003ZG03NP1&cluster=7624.f. Contact :
          Feng Liang Email : fliang@lifetech.com URL :
          http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
          Faraday Avenue Genoscope sequence ID : CS0AS003ZG03NP1.
          Location/Qualifiers
            1..1049
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="CS0DB005YL20"

```

```

FEATURES
source

```

```

/tissue type="NEUROBLASTOMA COT 10-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 58.0%; Score 705.2; DB 13; Length 1049;
Best Local Similarity 81.4%; Pred. No. 9.3e-152;
Matches 829; Conservative 36; Mismatches 141; Indels 13; Gaps 5;

171 GAATTTGAAACTTTAAACCTTAAATGGTATATACATCCAAAGAGACTATCTCTTAAGT 230
Db      1017 GAATTTGAACTTTGAACTTTAACTAAAYKGTCTTACCCACAMMMACTACTCTCTWAG 958
Qy      231 AAGAAATTTCCATATTCCTTTGGAACTCTTACAGACTCTTACTGTGGGCTTGTCCGAGTT 290
Db      957 TARGAATTTCCAAATYCTTGG---ACCTTCTCATCTCTWATTGKGGCTGTCCGKMTT 902
Qy      291 GATATGGTATGCTTTCCTTAAAGAGCTTAGGAAATAGACTTGTAGTCAACACATATA 350
Db      901 -ATACGGCTGTGTTCCTTTTAMA---SCCTTGGACCTTMACTTTAGTMAACAATATA 846
Qy      351 AAAAAAGCTTCCAGCTCAAAATTTGGAGACTCATACACCTTCAAGAACTTTAACTCGAATGAC 410
Db      845 AAAAAAG-TTCTCTGCTACAAATTTGGAGACTTTATACACCTTTAAGAACTTTAACTCGAATGAT 787
Qy      411 AATCAGTTGGAGTCAATTTAGTGTAGCCTTGTGTCTTACTTCTACACTCCAGAGTCACTTCGG 470
Db      786 AATMACTTGGAGTCAATTTAGTGTAGCCTTGTGTCTTACTTCTACACTCCAGAGTCACTTCGG 727
Qy      471 AGTTTGGACCTCAGCAGAAACAAATCAAGCACTCCCTGTCTGAGTTTGGCACTCCAG 530
Db      726 AGTTTGGACCTCAGCAGAAACAAATCAAGCACTCCCTGTCTGAGTTTGGCACTCCAG 667
Qy      531 GAATTAAGAAATTTAAACCTTGGAGATAATGAATTTGAATTTCTTCTGCGAAGATAGGA 590
Db      666 GAATTAAGAAATTTAAACCTTGGAGATAATGAATTTGAATTTCTTCTGCGAAGATAGGA 607
Qy      591 CAATTAATAAACCTTCGCTTTTGTTCAGAGCTCGAAATTAAGCTTCCATTTTTCCTAGT 650
Db      606 CAATTAATAAACCTTCGCTTTTGTTCAGAGCTCGAAATTAAGCTTCCATTTTTCCTAGT 547
Qy      651 GAATTTAGAAATTTATCCCTTGAATCTCTGAGTCTTTTGGAAATCTTTTGGACACCA 710
Db      546 GAATTTAGAAATTTATCCCTTGAATCTCTGAGTCTTTTGGAAATCTTTTGGACACCA 487
Qy      711 AAAGTCTCTTCCAGT---AATAAAGCTGCAAGCACCATTAACTTTATTTGAAATCTTCTGC 766
Db      486 AGAGTCTTTCGGGTGCTGTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 427
Qy      767 ACCAACCATTATACATAAATAGGATTCATATGGCTCTCATATTCATTCATTCATTCAT 826
Db      426 ACCAACCATTATACATAAATAGGATTCATATGGCTCTCATATTCATTCATTCATTCAT 367
Qy      827 CCAAGATTTGGATACCGCAAAATTTTGTTCGTGGAAGATTTCTGCTGAACCTTTTCAT 886
Db      366 CCAAGATTTGGATACCGCAAAATTTTGTTCGTGGAAGATTTCTGCTGAACCTTTTCAT 307
Qy      887 TCAAGAACTTACTTACCATGAATCTGATCTGTGTGCCCACACTGTGTGTCTTAGTAGATA 946
Db      306 TCAAGAACTTACTTACCATGAATCTGATCTGTGTGCCCACACTGTGTGTCTTAGTAGATA 247
Qy      947 TTTGGGTGTACTGAAGCACTTATATCTCTTATTTCTGTCTCTAGGCTGTATAGTTAA 1006
Db      246 TTTGGGTGTACTGAAGCACTTATATTTTATTTCTGTCTCTAGGCTGTATAGTTAA 187
Qy      1007 TTTCTCTGATATGTTAAAGTAATGGGTGAGACCCAGAAAGAAATTTCAATACAGATCA 1066
Db      186 TTTCTCTGATATGTTAAAGTAATGGGTGAGACCCAGAAAGAAATTTCAATACAGATCA 127
Qy      1067 GTTTGGGTGTCATGTATGATTTTTCGACGCTCAAAATTTGGAGTAGGGAAGATTTCTGTATA 1126

```

```

Db      126  GTTTGGGTCATGATGATTTTGACGCTCATATTCGAGTAAGGAGATTTTGTATA 67
QY      1127  CTTGCTGGAGAGAGAAATGTTATAGTACTCATTTAGTAGTACTCCAAACATTTTATT 1185
Db      66   TTTGTGGAGTGGTGAATGTTGTGTTGTTATTTATTTTGTGTTATTTTATTTTNTT 8

RESULT 13
BU620411/c
LOCUS   BU620411 722 bp mRNA linear EST 23-SEP-2002
DEFINITION
UI-R-FL1-bfw-1-09-0-UI.s1 NCI CGAP FL1 Homo sapiens cDNA clone
UI-H-FL1-bfw-1-09-0-UI 3', mRNA sequence.
ACCESSION
BU620411
VERSION
BU620411.1 GI:23286626
KEYWORDS
EST.
SOURCE  Homo sapiens (human)
ORGANISM
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
1 (bases 1 to 722)
AUTHORS
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
CONTACT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Martin
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@iowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-43, sAT rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
Location/Qualifiers
1..722
/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/clone="UI-H-FL1-bfw-1-09-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FL1"
/note="Organ: Chondrosarcoma; Vector: pT7T3-Pac
(Pharmacia) with a modified polylinker; Site 1: EcoR I;
Site 2: Not I; NCI CGAP FL1 is a normalized cDNA library
derived from a pool of mRNA obtained from 4 cell lines
from grade III chondrosarcoma tissues. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(drl18 tail). The sequence tag for this library is
GAGTCCGTTG. The cell lines were provided by Dr. James
Martin from the University of Iowa.
TAG TISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG LIB=UI-H-FL1
TAG_SEQ=GAGTCCGTTG"

ORIGIN
Query Match 57.9%; Score 763.8; DB 13; Length 722;
Best Local Similarity 99.0%; Pred. No. 2e-151;
Matches 708; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      501  GCATCCCTGTCAGTTTGGCAGCTCCAGGAACCTTAAGAAATTTAAACCTTGACGATAAT 560

```

```

Db      721  GCATCCCTGTCAGTTTGGCAGCTCCAGGAACCTTAAGAAATTTAAACCTTGACGATAAT 662
QY      561  GAATGAAATCAATTTCTTGCAGAGTAGGCAACTAATAAACCTTCCTTTTGTGACGA 620
Db      661  GAATGAAATCAATTTCTTGCAGAGTAGGCAACTAATAAACCTTCCTTTTGTGACGA 602
QY      621  GCTGGAATAAGCTTCCATTTTGGCTAGTGAATTTAGAAATTTATCCCTTGAATCTG 680
Db      601  GCTGGAATAAGCTTCCATTTTGGCTAGTGAATTTAGAAATTTATCCCTTGAATCTG 542
QY      681  GATCTTTTGGAAATATCTTTTGAACCAACCAAAAGTCTCTCCAGTAATAAAGCTCAAGCA 740
Db      541  GATCTTTTGGAAATATCTTTTGAACCAACCAAAAGTCTCTCCAGTAATAAAGCTCAAGCA 482
QY      741  CCATTAACCTTTATGGAATCTTCTGCAACCAATATTACATAATAGGATTCATATGGC 800
Db      481  CCATTAACCTTTATGGAATCTTCTGCAACCAATATTACATAATAGGATTCATATGGC 422
QY      801  TCTCATATCATCTCCATCTCTGCGAGATTTGGATACCCGCAAAATTTGTGTTGT 860
Db      421  TCTCATATCATCTCCATCTCTGCGAGATTTGGATACCCGCAAAATTTGTGTTGT 362
QY      861  GGAAGATTTCTGCTGGAACCTTTTCATTCAAGGAACCTACTACCATGAATCTGCTGTT 920
Db      361  GGAAGATTTCTGCTGGAACCTTTTCATTCAAGGAACCTACTACCATGAATCTGCTGTT 302
QY      921  GCCCACACTGTGCTTTAGTAGAATAATTTGGTGGTACTGAGACCTATTATCTCTAT 980
Db      301  GCCCACACTGTGCTTTAGTAGAATAATTTGGTGGTACTGAGACCTATTATCTCTAT 242
QY      981  TTTCTGTTCTTAGCGGTATGTTTATTTCTCTGATATGTTAAAGTAATGGGTGAGACCA 1040
Db      241  TTTCTGTTCTTAGCGGTATGTTTATTTCTCTGATATGTTAAAGTAATGGGTGAGACCA 182
QY      1041  GAAAAAGAAATTTCAATAACAGATCAGTTTGGGTGCTATGATGATTTTCAGCGTCAAA 1100
Db      181  GAAAAAGAAATTTCAATAACAGATCAGTTTGGGTGCTATGATGATTTTCAGCGTCAAA 122
QY      1101  TTGGAGTAAGGGAAGATTTCTGTATATCTTCTGAGAGGAGGAATGTGTATGTTACTCA 1160
Db      121  TTGGAGTAAGGGAAGATTTCTGTATATCTTCTGAGAGGAGGAATGTGTATGTTACTCA 62
QY      1161  TTTAGATGACTCCAAACCTTTTATTAAACCAATTTAGTTTAAATAAAAAA 1215
Db      61   TTTAGATGACTCCAAACCTTTTATTAAACCAATTTAGTTTAAATAAAAAA 7

RESULT 14
BX367719
LOCUS   BX367719
DEFINITION
BX367719 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
cDNA clone CSDB005YM21 5-PRIME, mRNA sequence.
ACCESSION
BX367719
VERSION
BX367719.1 GI:30437265
KEYWORDS
EST.
SOURCE  Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
1 (bases 1 to 931)
AUTHORS
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished (2001)
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSDB005Y2A07_T0433_1.
Location/Qualifiers

```

FEATURES


```
source 1. .891
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0B005YM21"
/tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 57.4%; Score 700; DB 13; Length 891;
Best Local Similarity 97.0%; Pred. NO. 1.5e-150;
Matches 733; Conservative 0; Mismatches 19; Indels 4; Gaps 2;
QY 460 AGTCACCTCGGAGTTTGGACCTCAGCAAGAACAAATCAAGGACCTCCCTGTGCACTTTT 519
DB 20 AGTCACCTCGGAGTTTGGACCT---CCCAAGACAAAATCAAGCACTCCCTGTGCACTTTT 76
QY 520 GCACGCTCAGGAACCTTAAGATTTAAACTTTGACGATATGAATGATTCATTTCCCTT 579
DB 77 GCACGCTCAGGAACCTTAAGATTTAAACTTTGACGATATGAATGATTCATTTCCCTT 136
QY 580 GCAAGATAGCAACCTTAATAAACCTTCGCTTTTGTGACGAGCTCGAAATTAAGCTTCCAT 639
DB 137 GCAGATAGGACACCTAATTAACCTTCGCTTTTGTGACGAGCTCGAATAAGCTTCCAT 196
QY 640 TTTTGCTAGTGAATTTAGAAATTTATCCCTTGAATACCTTGGATCTTTTGGAAATACCTT 699
DB 197 TTTTGCTAGTGAATTTAGAAATTTATCCCTTGAATACCTTGGATCTTTTGGAAATACCTT 256
QY 700 TTGAACAACCAAAAGCTCTCCAGTAAATTAAGCTGCAAGCACCATTAATTTATTCGAAT 759
DB 257 TTGAACAACCAAAAGCTCTCCAGTAAATTAAGCTGCAAGCACCATTAATTTATTCGAAT 316
QY 760 CTTCTGCAGCAACCATATTACATAATAGATTCATATGCTCTCATATCATTCATTC 819
DB 317 CTTCTGCAGCAACCATATTACATAATAGATTCATATGCTCTCATATTCATTC 376
QY 820 ATCTCTGCCAAGATTTGGATACCGCAAAAATTTGTGTTTGTGGAAGATCTGTCTGAACCT 879
DB 377 ATCTCTGCCAAGATTTGGATACCGCAAAAATTTGTGTTTGTGGAAGATCTGTCTGAACCT 436
QY 880 CTTTCATTCAGGAACCTACTACCATGAATCTGATTCGTGTCGCCACACTGTGCTTAG 939
DB 437 CTTTCATTCAGGAACCTACTACCATGAATCTGATTCGTGTCGCCACACTGTGCTTAG 496
QY 940 TAGATAAATTTGGTGGTACTGGAAGCACTTATATCTCTTATTTCTGTCTTAGGCTGTT 999
DB 497 TAGATAAATTTGGTGGTACTGGAAGCACTTATATCTCTTATTTCTGTCTTAGGCTGTT 556
QY 1000 ATGTTAAATTCCTCTGATGTTTAAAGTAAATCGGTGAGACAGAAAAGAAATTTCAATAA 1059
DB 557 ATGTTAAATTCCTCTGATGTTTAAAGTAAATCGGTGAGACAGAAAAGAAATTTCAATAA 616
QY 1060 CAGATCAGTTTGGGGTGCATGATGATTTTGGACGCTCAATTTGGAGTAAAGGAAGATTT 1119
DB 617 CAGATCAGTTTGGGGTGCATGATGATTTTGGACGCTCAATTTGGAGTAAAGGAAGATTT 676
QY 1120 CTGTATATCTGTGGAGAGAGAAATGTGTATAGTTACTCAATTTAGATGATCTCAAAACT 1179
DB 677 CTGTATATCTGTGGAGAGGA-GAAATGTGTATAGTACTCAATTTAGATGATCTCAAAACT 735
QY 1180 TTATTTAAACCAATTTTAGTTTAAATAAAAAA 1215
DB 736 TTATTTAAACCAATTTTAGTTTAAATAAAAAA 771

RESULT 15
BX394855/c
LOCUS BX394855 1165 bp mRNA linear EST 13-MAY-2003
```

```
DEFINITION BX394855 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DC024YC05 3-PRIME, mRNA sequence.
ACCESSION BX394855
VERSION BX394855.1 GI:30628345
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1165)
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7624.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DC024AB03NP1&cluster=7624.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DC024AB03NP1.
FEATURES
Location/Qualifiers
1..1165
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC024YC05"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 57.3%; Score 695.8; DB 13; Length 1165;
Best Local Similarity 97.8%; Pred. NO. 1.3e-149;
Matches 720; Conservative 6; Mismatches 10; Indels 2; Gaps 2;
QY 454 TCAGAAAGTCACTTCGGAGTTTGGACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGC 513
DB 741 TCAGAAAGTCACTTCGGAGTTTGGACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGC 682
QY 514 AGTTTGGCCAGCTCCAGGAACCTTAAGATTTAAACTTTGACGATATGAATTCATTCAT 573
DB 681 AGTTTGGCCAGCTCCAGGAACCTTAAGATTTAAACTTTGACGATATGAATTCATTCAT 622
QY 574 TTCCTTGCAGATAGGACCACTAATAAACCTTCGCTTTTGTGACGAGCTCGAAATAGC 633
DB 621 TTCCTTGCAGATAGGACCACTAATAAACCTTCGCTTTTGTGACGAGCTCGAAATAGC 562
QY 634 TTCATTTTTCCTAGTGAATTTAGAAATTTATCCCTTGAATACCTTGGATCTTTTGGAA 693
DB 561 TTCATTTTTCCTAGTGAATTTAGAAATTTATCCCTTGAATACCTTGGATCTTTTGGAA 502
QY 694 ATACTTTTGAACCAAGAGTCTTCAGTAAATAAGCTGCAAGCAACCAATTAACCTTAT 753
DB 501 ATACTTTTGAACCAAGAGTCTTCAGTAAATAAGCTGCAAGCAACCAATTAACCTTAT 442
QY 754 TGGAAATCTCTGCACCAACCATATTACATAATAGGATTCATATGCTCTCATATCATTC 813
DB 441 TGGAAATCTCTGCACCAACCATATTACATAATAGGATTCATATGCTCTCATATCATTC 382
QY 814 CATTCATCTCTGCCAAGATTTGGATACCGCAAAAATTTGTGTTTGTGGAAGATTCCTGC 873
DB 381 CATTCATCTCTGCCAAGATTTGGATACCGCAAAAATTTGTGTTTGTGGAAGATTCCTGC 322
QY 874 TGAATCTTTTCAATTAAGGAACCTACTACCATGAAATCTGCAATTCCTGTTGCCACACTGTGG 933
```


Db 321 TGAACCTCTTCATTCAAGGAACACTACCATGAATCTGCAATTCGTGTCGCCACACTGTGG 262
Qy 934 TCCTTAGTAGATAAATTTGGGTGGTACTGAGCACCTATTATCTCTTATTCTCTGTTCTCTAG 993
Db 261 TCTTAGTAGATAAATTTGGGTGGTACTGAGCACCTATTATCTCTTATTCTCTGTTCTCTAG 202
Qy 994 GCTGTTATGTTAAATTCCTCTGATATGTTAAAGTAATGGTGAGACCAAGAAAAGAAATTT 1053
Db 201 SCTGTTATGTTAAATTCCTCTGATATGTTAAAGTAATGGTGAGACCAAGAAAAGAAATTT 142
Qy 1054 CAATAACAGATCAGTTTGGGTGGCAATGATGATTTTGCAGGTCMAATGGAGTAAGGGA 1113
Db 141 CAATAACAGATCAGTTTGGGTGGCAATGATGATTTTGCAGGTCMAATGGAGTAAGGGA 82
Qy 1114 AGATTTCTGTATACTTGTGAGAGGAGGAATGTGTA-TAGTTACTC-ATTAGATGACT 1171
Db 81 AGATTTCTGTATACTTGTGAGAGGAGGAATGTGTA-TAGTTACTC-ATTAGATGACT 22
Qy 1172 CCAAAACCTTTTATTAAAA 1189
Db 21 TTAANNNTWHDARD 4

Search completed: June 16, 2004, 00:57:25
Job time : 2403 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2004, 23:13:58 ; Search time 418 seconds
(without alignments)
13267.648 Million cell updates/sec

Title: US-10-009-557-34

Perfect score: 1215

Sequence: 1 gaagaactagcatgtatga.....ttagtttttaaaaaaaaaa 1215

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2998549 seqs, 2282253817 residues

Total number of hits satisfying chosen parameters: 5997038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgm2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgm2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 3: /cgm2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:*
- 4: /cgm2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgm2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgm2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 7: /cgm2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgm2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgm2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgm2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgm2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgm2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgm2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 14: /cgm2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgm2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgm2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgm2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgm2_6/ptodata/1/pubpna/US62_NEW_PUB.seq:*
- 19: /cgm2_6/ptodata/1/pubpna/US62_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1153.2	94.9	1471	15	US-10-037-270-1099
2	1153.2	94.9	1471	16	US-10-117-722-1099
3	950.2	78.2	1251	16	US-10-312-088-22
4	777	64.0	780	16	US-10-312-088-21
5	715.8	58.9	1241	9	US-09-833-381-1401
6	521.6	42.9	682	9	US-09-813-358-33
7	521.6	42.9	682	10	US-09-937-279-33
8	344	28.3	487	13	US-10-085-783A-7737
9	344	28.3	487	16	US-10-242-535A-7737
10	302.8	24.9	564	15	US-10-029-386-7642
11	281.4	23.2	732	13	US-10-076-555-463
12	281	23.1	732	13	US-10-076-555-467
13	239	19.7	239	15	US-10-029-386-21342
14	87.2	7.2	582	9	US-09-778-927A-28

15	74.8	6.2	172	13	US-10-085-783A-5829
16	74.8	6.2	172	16	US-10-242-535A-5829
17	60.2	5.0	2950	13	US-10-220-120-21
18	55	4.5	693	16	US-10-369-493-34864
19	55	4.5	4984	17	US-10-468-334-15
20	50	4.1	50	16	US-10-131-827-7581
21	49.8	4.1	2413	13	US-10-342-887-1742
22	49.8	4.1	2413	13	US-10-172-118-1742
23	48.2	4.0	1947	9	US-09-909-320-184
24	48.2	4.0	1947	9	US-09-909-088B-184
25	48.2	4.0	1947	9	US-09-905-291A-184
26	48.2	4.0	1947	9	US-09-902-853-184
27	48.2	4.0	1947	9	US-09-907-824-184
28	48.2	4.0	1947	9	US-09-907-841-184
29	48.2	4.0	1947	10	US-09-904-011-184
30	48.2	4.0	1947	10	US-09-906-742-184
31	48.2	4.0	1947	10	US-09-906-838-184
32	48.2	4.0	1947	10	US-09-907-613-184
33	48.2	4.0	1947	10	US-09-907-942-184
34	48.2	4.0	1947	10	US-09-904-859-184
35	48.2	4.0	1947	10	US-09-908-204-184
36	48.2	4.0	1947	10	US-09-904-820-184
37	48.2	4.0	1947	10	US-09-904-786-184
38	48.2	4.0	1947	10	US-09-906-646-184
39	48.2	4.0	1947	10	US-09-906-700-184
40	48.2	4.0	1947	10	US-09-903-786-184
41	48.2	4.0	1947	10	US-09-902-903-184
42	48.2	4.0	1947	10	US-09-903-749A-184
43	48.2	4.0	1947	10	US-09-904-119-184
44	48.2	4.0	1947	10	US-09-904-956-184
45	48.2	4.0	1947	10	US-09-902-736-184

ALIGNMENTS

RESULT 1
US-10-037-270-1099
; Sequence 1099, Application US/10037270
; Publication NO. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-Hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE INVENTION: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/486,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: Pf-FL_genes Version 1.0
; SEQ ID NO 1099
; LENGTH: 1471
; TYPE: DNA
; ORGANISM: Homo sapiens

```

FEATURE:
NAME/KEY: CDS
LOCATION: (40)..(1284)
US-10-037-270-1099

Query Match
Best Local Similarity 94.9%; Score 1153.2; DB 15; Length 1471;
Matches 1155; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 58 TTATTCAGGCCAATTCACAGCTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCATTA 117
DB 314 TAAAGTAAGGCCAATTCACAGCTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCATTA 373
QY 118 GAGGCTGTAATGTTGATACACCACTTCACAGCTTCACAGCTTCACAGCTTCAGAAATTCG 177
DB 374 GAGGCTGTAATGTTGATACACCACTTCACAGCTTCACAGCTTCACAGCTTCAGAAATTCG 433
QY 178 AAAAATTTAAACTAAATAGGTTTATCACATCCCAAAAAGACTATCCTCTAAGTAAGAATT 237
DB 434 AAAAATTTAAACTAAATAGGTTTATCACATCCCAAAAAGACTATCCTCTAAGTAAGAATT 493
QY 238 TTCCAGTAATTAAGCTGCAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 957
DB 1094 ATACCGCAAAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1153
QY 898 CTACCAATGAATCTGCAATTCCTGTTGCCACACTGGTCTTAGTAGATAATTTGGTGGTA 957
DB 1154 CTACCAATGAATCTGCAATTCCTGTTGCCACACTGGTCTTAGTAGATAATTTGGTGGTA 1213
QY 958 CTGAAGCACTATTAATCTCTATTTCTGTTCTCTAGCTGTTATGTTAATTCCTCTGATA 1017
DB 1214 CTGAAGCACTATTAATCTCTATTTCTGTTCTCTAGCTGTTATGTTAATTCCTCTGATA 1273

RESULT 2
US-10-117-722-1099
: Sequence 1099, Application US/10117722
: Publication No. US20030219744A1
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Liu, Chenghua
: APPLICANT: Asundi, Vinod
: APPLICANT: Zhang, Jie
: APPLICANT: Drmanac, Radoje T.
: TITLE OF INVENTION: No. US20030219744A1elel Nucleic Acids and
: TITLE OF INVENTION: Polypeptides
: FILE REFERENCE: 784CIP2BCIP
: CURRENT APPLICATION NUMBER: US/10/117,722
: CURRENT FILING DATE: 2002-04-04
: PRIOR APPLICATION NUMBER: 09/620,312
: PRIOR FILING DATE: 2000-07-19
: PRIOR APPLICATION NUMBER: 09/552,317
: PRIOR FILING DATE: 2000-04-25
: PRIOR APPLICATION NUMBER: 09/488,725
: PRIOR FILING DATE: 2000-01-21
: NUMBER OF SEQ ID NOS: 1104
: SOFTWARE: PF_FL_genes Version 1.0
: SEQ ID NO 1399
: LENGTH: 1471
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (40)..(1284)
US-10-117-722-1099

Query Match
Best Local Similarity 94.9%; Score 1153.2; DB 16; Length 1471;
Matches 1155; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 58 TTATTCAGGCCAATTCACAGCTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCATTA 117
DB 314 TAAAGTAAGGCCAATTCACAGCTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCATTA 373
QY 118 GAGGCTGTAATGTTGATACACCACTTCACAGCTTCACAGCTTCACAGCTTCAGAAATTCG 177
DB 374 GAGGCTGTAATGTTGATACACCACTTCACAGCTTCACAGCTTCACAGCTTCAGAAATTCG 433
QY 178 AAAAATTTAAACTAAATAGGTTTATCACATCCCAAAAAGACTATCCTCTAAGTAAGAATT 237
DB 434 AAAAATTTAAACTAAATAGGTTTATCACATCCCAAAAAGACTATCCTCTAAGTAAGAATT 493
QY 238 TTCCAGTAATTAAGCTGCAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 957
DB 1094 ATACCGCAAAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1153
QY 898 CTACCAATGAATCTGCAATTCCTGTTGCCACACTGGTCTTAGTAGATAATTTGGTGGTA 957
DB 1154 CTACCAATGAATCTGCAATTCCTGTTGCCACACTGGTCTTAGTAGATAATTTGGTGGTA 1213
QY 958 CTGAAGCACTATTAATCTCTATTTCTGTTCTCTAGCTGTTATGTTAATTCCTCTGATA 1017
DB 1214 CTGAAGCACTATTAATCTCTATTTCTGTTCTCTAGCTGTTATGTTAATTCCTCTGATA 1273

```

```
QY 358 TTCCAGCTACATTGAGAGCTCATACACCTTCAAGAACTTACCTGATGCAATCACT 417
Db 614 TTCCAGCTACATTGAGAGCTCATACACCTTCAAGAACTTACCTGATGCAATCACT 673
QY 418 TGGAGTCATTAGTGTAGCTTGTGTCATCTACACTCCAGAGTCACTTCGGAGTTGG 477
Db 674 TGGAGTCATTAGTGTAGCTTGTGTCATCTACACTCCAGAGTCACTTCGGAGTTGG 733
QY 478 ACTCAGCAAGAACAAAATCAAGGCACTCCCTGTGCGAGTTTTCGCAAGCTCCAGGAACCTTA 537
Db 734 ACTCAGCAAGAACAAAATCAAGGCACTCCCTGTGCGAGTTTTCGCAAGCTCCAGGAACCTTA 793
QY 538 AGAATTTAAACCTTGACGATGAATGAATGATTCATTTCTTCCAGAGATGAGCAACTTA 597
Db 794 AGAATTTAAACCTTGACGATGAATGAATGATTCATTTCTTCCAGAGATGAGCAACTTA 853
QY 598 TAAACCTTCGCTTTTGTGAGAGCTCGAAATAAGCTTCCATTTTTCGCTAGTGAATTTA 657
Db 854 TAAACCTTCGCTTTTGTGAGAGCTCGAAATAAGCTTCCATTTTTCGCTAGTGAATTTA 913
QY 658 GAAATTTATCCCTTGATATCTTGGATCTTTTGGAAATACCTTTGAAACACCAAAAGTCC 717
Db 914 GAAATTTATCCCTTGATATCTTGGATCTTTTGGAAATACCTTTGAAACACCAAAAGTCC 973
QY 718 TTCCAGTAATAAGCTGCAAGCACCAATTAACCTTATTGGAATCTTCTGCAAGAACCATAT 777
Db 974 TTCCAGTAATAAGCTGCAAGCACCAATTAACCTTATTGGAATCTTCTGCAAGAACCATAT 1033
QY 778 TACATAATAGGAATCCATATGCTCTCATATCATTCATCCATTCCTGCAAGATTTGG 837
Db 1034 TACATAATAGGAATCCATATGCTCTCATATCATTCATCCATTCCTGCAAGATTTGG 1093
QY 838 ATACCGCAAAATTTGTGTTTGGAGAGATTCTGCTGAACTCTTTCATTCAGGAACTA 897
Db 1094 ATACCGCAAAATTTGTGTTTGGAGAGATTCTGCTGAACTCTTTCATTCAGGAACTA 1153
QY 898 CTACCATGAATCTGCATCTGTTGTCACACCTGCTGCTTAGTAGATAATTTGGGTGGTA 957
Db 1154 CTACCATGAATCTGCATCTGTTGTCACACCTGCTGCTTAGTAGATAATTTGGGTGGTA 1213
QY 958 CTGAAGCACTTAATATCTTATTTCTGTTCTCTAGGCTGTTATGTTAATCTCTCGATA 1017
Db 1214 CTGAAGCACTTAATATCTTATTTCTGTTCTCTAGGCTGTTATGTTAATCTCTCGATA 1273
QY 1018 TGTAAAGTAATGGTGGAGACAGAAAAGAAATTTCAATAACAGATCAAGTTTGGGGTGC 1077
Db 1274 TGTAAAGTAATGGTGGAGACAGAAAAGAAATTTCAATAACAGATCAAGTTTGGGGTGC 1333
QY 1078 ATGTATGATTTTGCAGCTGCAAAATGGAGTAAGGGAAGATTTCTGTATATCTTGTGAGGA 1137
Db 1334 ATGTATGATTTTGCAGCTGCAAAATGGAGTAAGGGAAGATTTCTGTATATCTTGTGAGGA 1393
QY 1138 GGAGGAATCTATATGTTACTCATTTAGATGACTCCAAACTTTTATTAACCAATTTT 1197
Db 1394 GGAGGAATCTATATGTTACTCATTTAGATGACTCCAAACTTTTATTAACCAATTTT 1453
QY 1198 AGTTTAAAAAATAAAAA 1215
Db 1454 AGTTTAAAAAATAAAAA 1471
```

RESULT 3

US-10-312-088-22

; Sequence 22, Application US/10312088

; Publication No. US20030219862A1

; GENERAL INFORMATION:

; APPLICANT: Agarwal, Pankaj

; APPLICANT: Cogswell, John P.

; APPLICANT: Kabnic, Karen S.

; APPLICANT: Lai, Ying-Ta

; APPLICANT: Martensen, Shelby A.

; APPLICANT: Murdoch, Paul R.

; APPLICANT: Smith, Randall F.

```
; APPLICANT: Strum, Jay C.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Xie, Qing
; APPLICANT: Rizni, Safia K.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50029
; CURRENT APPLICATION NUMBER: US/10/312,088
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: PCT/US01/19929
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,161
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/213,156
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-312-088-22
```

Query Match

78.2%; Score 950.2; DB 16; Length 1251;

Best Local Similarity 99.1%; Pred. No. 3.7e-234;

Matches 968; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

```
QY 58 TTATTCAGCGCCATTTCAGAGTTTAAAGTTTCCCTTCAGCTATGAGACTGGCTCAT 117
Db 275 TAAGTAAGGCCATTTCAGAGTTTAAAGTTTCCCTTCAGCTATGAGACTGGCTCAT 334
QY 118 GAGGCTGTAATGTGTATACACAGTTTCAAGCTCACACAGTGAAGACTTCAGAAATTG 177
Db 335 GAGGCTGTAATGTGTATACACAGTTTCAAGCTCACACAGTGAAGACTTCAGAAATTG 394
QY 178 AAAACCTTTAAAACTAAATGGTTATCACATCCAAAAAGACTATCTCTAAGTAAAGATT 237
Db 395 AAAACCTTTAAACTAAATGGTTATCACATCCAAAAAGACTATCTCTAAGTAAAGATT 454
QY 238 TTCCATATTCCTTGGAACTCTTTCAGACTTCTTCTGTGGGCTTGTCCGAGTTGATAAGC 297
Db 455 TTCCATATTCCTTGGAACTCTTTCAGACTTCTTCTGTGGGCTTGTCCGAGTTGATAAGC 514
QY 298 GTATCGTTTGTCTTAAAAAGCCCTTAGGAAATTAGAGTTAGAGTCAACACCATATAAAAAAGC 357
Db 515 GTATCGTTTGTCTTAAAAAGCCCTTAGGAAATTAGAGTTAGAGTCAACACCATATAAAAAAGC 574
QY 358 TTCCAGCTACAAATGGAGACCTCATACACCTTCAAGAACTTAACCTGAATGACAACTCACT 417
Db 575 TTCCAGCTACAAATGGAGACCTCATACACCTTCAAGAACTTAACCTGAATGACAACTCACT 634
QY 418 TGGAGTCATTAGTGTAGCTTGTGTCATCTACACTCCAGAGTCACTTCGGAGTTGG 477
Db 635 TGGAGTCATTAGTGTAGCTTGTGTCATCTACACTCCAGAGTCACTTCGGAGTTGG 694
QY 478 ACTCAGCAAGAACAAAATCAAGGCACTCCCTGTGCGAGTTTTCGCAAGCTCCAGGAACCTTA 537
Db 695 ACTCAGCAAGAACAAAATCAAGGCACTCCCTGTGCGAGTTTTCGCAAGCTCCAGGAACCTTA 754
QY 538 AGAATTTAAACCTTGACGATGAATGAATGATTCATTTCTTCCAGAGATGAGCAACTTA 597
Db 755 AGAATTTAAACCTTGACGATGAATGAATGATTCATTTCTTCCAGAGATGAGCAACTTA 814
QY 598 TAAACCTTCGCTTTTGTGAGAGCTCGAAATAAGCTTCCATTTTTCGCTAGTGAATTTA 657
Db 815 TAAACCTTCGCTTTTGTGAGAGCTCGAAATAAGCTTCCATTTTTCGCTAGTGAATTTA 874
QY 658 GAAATTTATCCCTTGATATCTTGGATCTTTTGGAAATACCTTTTGAACACCAAAAGTCC 717
Db 875 GAAATTTATCCCTTGATATCTTGGATCTTTTGGAAATACCTTTTGAACACCAAAAGTCC 934
QY 718 TTCCAGTAATAAGCTGCAAGCACCAATTAACCTTATTGGAATCTTCTGCAAGAACCATAT 777
Db 935 TTCCAGTAATAAGCTGCAAGCACCAATTAACCTTATTGGAATCTTCTGCAAGAACCATAT 994
```



```
Db 491 AAAACCTTAAACCTAAATGCTTATCATCCAAAAGAGCTATCTCTAAGTAAGATT 550
Qy 238 TTCCATATCTCTTGGAAACATCTTCAGACTTCTTACTGTGTGGCTGTGCGAGTTGATATGC 297
Db 551 TTCCATATCTCTTGGAAACATCTTCAGACTTCTTACTGTGTGGCTGTGCGAGTTGATATGC 610
Qy 298 GTATGCTTTGCTTAAAGAGCTTAGGAAATAGACTTGTAGTCTCAACCATATATAAAGC 357
Db 611 GTAIGCTTTGCTTAAAGAGCTTAGGAAATAGACTTGTAGTCTCAACCATATATAAAGC 670
Qy 358 TTCCAGCTCAATTTGGAGAGCTCATACRCCTTCAAGAACTTAACTGAAATGACATCACT 417
Db 671 TTCCAGCTCAATTTGGAGAGCTCATACRCCTTCAAGAACTTAACTGAAATGACATCACT 730
Qy 418 TGGAGTCATTTAGTGTAGCTTGTGTCATCTTACCTTCACTCCAGAGTGCATTTGGAGTTGG 477
Db 731 TGGAGTCATTTAGTGTAGCTTGTGTCATCTTACCTTCACTCCAGAGTGCATTTGGAGTTGG 790
Qy 478 ACCTCAGCAAGCAAAATCAAGGCACTCCCTGTGCAAGTTTTCGCACTCCAGGAACTTA 537
Db 791 ACCTCAGCAAGCAAAATCAAGGCACTCCCTGTGCAAGTTTTCGCACTCCAGGAACTTA 850
Qy 538 AGAATTTAAACCTTGA-CGATAATGAATTAATCAATTTCCCTTGAAGATAGGACAACTA 596
Db 851 AGAATTTAAACCTTGAACCGATATGAATTAATCAATTTCCCTTGAAGATAGGACAACTA 910
Qy 597 ATAAACCTTGCCTTTTGTGAGCAGCTCGAAATAGCTTCCATTTTGCCTAGTGAATTT 656
Db 911 ATAAACCTTGCCTTTTGTGAGCAGCTCGAAATAGCTTCCATTTTGCCTAGTGAATTT 970
Qy 657 AGAATTTATCCCTTGAATTAATGAATTTTGTGAAATTAATTTGAAACCAACCAAGTC 716
Db 971 AGAATTTATCCCTTGAATTAATGAATTTTGTGAAATTAATTTGAAACCAACCAAGTC 1030
Qy 717 CTTCAGTAAATAAGCTCAAGCAACCAATTAATTTTGAATTTCTTCGCAAGCAACATA 776
Db 1031 CTTCAGTAAATAAGCTCAAGCAACCAATTAATTTTGAATTTCTTCGCAAGCAACATA 1090
Qy 777 TTACATAATAGGATTCATATGCTCTCATATCAT 811
Db 1091 TTACATAATAGGATTAAGATTTTAAATAGTCAATGTAAT 1125
```

```
RESULT 6
US-09-813-358-33
; Sequence 33, Application US/09813358
; Patent No. US20020048759A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Pyle, Ruth
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.501
; CURRENT APPLICATION NUMBER: US/09/813,358
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 222
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(682)
; OTHER INFORMATION: n = A,T,C or G
US-09-813-358-33
```

Query Match 42.9%; Score 521.6; DB 9; Length 682;
Best Local Similarity 97.5%; Pred. No. 7.2e-124;
Matches 551; Conservative 0; Mismatches 10; Indels 4; Gaps 2;

```
Qy 277 GSCCTGTCCGAGTTGATGCTATGCTTAAAGAGCTTAGGAAATTAGACTTGA 336
Db 1 GSCCTGTCCGAGTTGATGCTATGCTTAAAGAGCTTAGGAAATTAGACTTGA 60
Qy 337 GTCACAAACCATATAAAGAGCTTCCAGCTACAAATGGAGAGCTCATACACCTTCAAGAAC 396
Db 61 GTCACAAACCATATAAAGAGCTTCCAGCTACAAATGGAGAGCTCATACACCTTCAAGAAC 420
Qy 397 TTAACCTGAATGCAATCACTTGGAGTCATTTAGTGTAGCTTGTGTCAATTCACATCC 456
Db 121 TTAACCTGAATGCAATCACTTGGAGTCATTTAGTGTAGCTTGTGTCAATTCACATCC 480
Qy 457 AGAAGTCACCTTCGAGTTTGGAGCTCAGCAAGCAAAATCAAGGCACTCCCTGTGAGT 516
Db 181 AGAAGTCACCTTCGAGTTTGGAGCTCAGCAAGCAAAATCAAGGCACTCCCTGTGAGT 540
Qy 517 TTTGCCAGCTCCAGGAACTTAAAGAAATTTAAACCTTGAAGATTAATGATTCATTTTC 576
Db 241 TTTGCCAGCTCCAGGAACTTAAAGAAATTTAAACCTTGAAGATTAATGATTCATTTTC 600
Qy 577 CTTGCAAGATAGGACAACTAATAAAGCTTGTGCTTGTGAGGAGCTCGAAATAGCTTC 636
Db 301 CTTGCAAGATAGGACAACTAATAAAGCTTGTGCTTGTGAGGAGCTCGAAATAGCTTC 660
Qy 637 CATTTTGGCTAGTGAATTTAGAAATTTATCCCTTGAATTAATGATTCATTTTGAATA 696
Db 361 CATTTTGGCTAGTGAATTTAGAAATTTATCCCTTGAATTAATGATTCATTTTGAATA 720
Qy 697 CTTTGAACCAACCAAGCTTCCAGTAAATAAGCTGCAAGCAACCAATTAATTTATTTGG 756
Db 421 CTTTGAACCAACCAAGCTTCCAGTAAATAAGCTGCAAGCAACCAATTTATTTATTTGG 780
Qy 757 AATCTCTGCAAGCAACCAATTTATTAATAGCTTGAATTCATATGCTCTCTCTCTCTCT 812
Db 481 AATCTCTGCAAGCAACCAATTTATTAATAGCTTGAATTCATATGCTCTCTCTCTCTCT 840
Qy 813 CCATTCATCTCTGCCAAGATTTGG 837
Db 541 CCATTCATCTCTGCCAAGATTTGG 865
```

```
RESULT 7
US-09-997-279-33
; Sequence 33, Application US/09997279
; Publication No. US20030059781A1
; GENERAL INFORMATION:
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN AND ENDOMETRIAL CANCER
; FILE REFERENCE: 210121.501C1
; CURRENT APPLICATION NUMBER: US/09/997,279
; CURRENT FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(682)
; OTHER INFORMATION: n = A,T,C or G
US-09-997-279-33
```

Query Match 42.9%; Score 521.6; DB 10; Length 682;
Best Local Similarity 97.5%; Pred. No. 7.2e-124;
Matches 551; Conservative 0; Mismatches 10; Indels 4; Gaps 2;

Qy 277 GSCCTGTCCGAGTTGATGCTATGCTTAAAGAGCTTAGGAAATTAGACTTGA 336
Db 1 GSCCTGTCCGAGTTGATGCTATGCTTAAAGAGCTTAGGAAATTAGACTTGA 60

RESULT 11
US-10-076-555-463
; Sequence 463, Application US/10076555
; Publication No. US20030065156A1
GENERAL INFORMATION:

Query Match	23.2%;	Score 281.4;	DB 13;	Length 750;
Best Local Similarity	87.0%;	Pred. No. 6e-62;		
Matches 320;	Conservative 0;	Mismatches 45;	Indels 3;	Gaps 2;
QY	786	AGGATTCGATATGGCTCTCATATCATTCATTCATTCCTGCCAAGATTGGATACGGCA	845	
Db	383	AGGATTCGATATGGCTCTCATATCATTCATTCATTCCTGCCAAGATTGGATACGGCA	442	
QY	846	AAAATTGTGTTTGTGGAAGATTCTGTCGAACTCTTTCATTCGAAGAACTACTACCATG	905	
Db	443	AAAATTGTGTGTNGGAAGATTCTGTCGAACTCTTTCATTCGAAGAACTACTACCATG	502	
QY	906	AATCTGCATTCTGTTGCCACACTGTGGTCTTAGTAGATAAATTTGGGTGGTACTGAAGCA	965	
Db	503	AATCTGCATTCTGNTGCCACACTGAGGNCCTTAGTAGATAAATTTGGGTGGTCTGAAGCA	560	
QY	966	CGTATTATCTCTAATTTCTGTTCTCTPAGGCTGTTATGTTTAATTCCTCTCATATGTT-AAA	1024	
Db	561	CCTATTATCTCTAATTTCTGTTCTCTPAGGCTGTTATGTTTAATTCCTCTGAATGNTAAAA	620	
QY	1025	GTAATGGGTGAGACCAAGAAAAAGAAATTCCTAATACAGATCAGTTTGGGGTGCATGTCATG	1084	
Db	621	GTAATGGGTGAGACCAAGAAAAAGAAATTCCTAATACAGATCAGTTTGGGGTGCATGTCATG	680	

QY 1085 ATTTTGCACGCTCAATTTGGAGTGAAGGAGATTTCTGTATATCTGTCTGAGAGGAGGAA 1144
DB 681 ATTTTCAAGCGTCAAAATGGAATAGGGAAGANTCTGGATACCTGCTTTGGAAAGGAAG 740
QY 1145 TGTGTATA 1152
DB 741 NATGTGTA 748

RESULT 12

US-10-076-555-467
; Sequence 467, Application US/10076555
; Publication No. US20030065156A1

; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pott, David
; APPLICANT: Kassam, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickinson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, Lee William
; APPLICANT: Stache-Crain, Birgit
; TITLE OF INVENTION: Novel Human Genes and Gene Expression
; FILE REFERENCE: 2300-1480
; CURRENT APPLICATION NUMBER: US/10/076,555
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/080,664
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/105,877
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-27
; NUMBER OF SEQ ID NOS: 844
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 467
; LENGTH: 732
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(732)
; OTHER INFORMATION: n = A,T,C or G

US-10-076-555-467
Query Match 23.1%; Score 281; DB 13; Length 732;
Best Local Similarity 90.9%; Pred. No. 7.5e-62;
Matches 329; Conservative 0; Mismatches 30; Indels 3; Gaps 3;

QY 786 AGAATTCATATGGCTCTCATATCCATTCATTCCTGCAAGATTTGGATACCGCA 845
DB 373 AGAATTCATATGGCTCTCATATCCATTCATTCCTGCAAGATTTGGATACCGCA 432
QY 846 AAAATTTGTTGTGGAGATTTCTGTGAACTTTTCAATTCAGGAACCTACTACATG 905

DB 433 AAAATTTGTTGTGGAGATTTCTGTGAACTTTTCAATTCAGGAACCTACTACATG 492
QY 906 AATCTGCATTTCTGTTGCCACACTGTGGTCTTAGTAGATAAATTTGGTGTACTGAGCA 965
DB 493 AATCTGCATTTCTGNTGCCACACTGTGGTCTTAGTAGATAAATTTGGTGTACTGAGCA 551
QY 966 CTTATTATCTCTTATTTCGTGTTCTTAGGCTGTATTGTTAAATTCCTCTGATATGTTAAAG 1025
DB 552 CTTATTATCTCTTATTTCGTGTTCTTAGGCTGTATTGTTAAATTCCTCTGATATGTTAAAG 611
QY 1026 TAATGGTGCAGACACAGAAAAGAAATTTCAATAACAGATCAAGTTTGGGGTGCATGATGA 1085
DB 612 TAATGGTGCAGACACAGAAAAGAAATTTCAATAACAGATCAAGTTTGGGGTGCATGATGA 670
QY 1086 TTTTGCAGGCTCAATTTGGAGTAAGGAGATTTCTGTATATCTGTCTGAGAGGAGGAAT 1145
DB 671 ATTTGCAACTCAATTTGGAGTAAGGAGGAANA-TTCTGGATACCTGCTGGAAGGAGGAAT 729
QY 1146 GT 1147
DB 730 GT 731

RESULT 13

US-10-029-386-21342
; Sequence 21342, Application US/10029386
; Publication No. US20030194704A1

; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G;
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21342
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR14.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EST HUMAN HIT: BG940337.1, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q92610, EVALUATE 1.00e+00
; OTHER INFORMATION: NT HIT: g114277125, EVALUATE 0.00e+00
US-10-029-386-21342

Query Match 19.7%; Score 239; DB 15; Length 239;
Best Local Similarity 100.0%; Pred. No. 2.7e-51;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 788 GATTCATATGGCTCTCATATCCATTCATTCCTGCAAGATTTGGATACCGCAA 847
DB 1 GATTCATATGGCTCTCATATCCATTCATTCCTGCAAGATTTGGATACCGCAA 60
QY 848 AATTTGTTTGTGGAGATTTCTGTGAACCTTTTCATTCAGGAACCTACTACATGAA 907
DB 61 AATTTGTTTGTGGAGATTTCTGTGAACCTTTTCATTCAGGAACCTACTACATGAA 120
QY 908 TCTGCATTTCTGTTGCCACACTGTGGTCTTAGTAGATAAATTTGGGGTGGTACTGAGCAC 967
DB 121 TCTGCATTTCTGTTGCCACACTGTGGTCTTAGTAGATAAATTTGGGGTGGTACTGAGCAC 180
QY 968 TATTACTCTTATTCTGTTCTTAGGCTGTATTGTTAAATTCCTCTGATATGTTAAAGT 1026
DB 181 TATTACTCTTATTCTGTTCTTAGGCTGTATTGTTAAATTCCTCTGATATGTTAAAGT 239

RESULT 14

US-09-778-927A-28
; Sequence 28, Application US/09778927A
; Patent No. US20020068342A1
; GENERAL INFORMATION:
; APPLICANT: KHOSRAVI, Rami et al.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID AND AMINO ACID SEQUENCES AND NOVEL
; FILE OF INVENTION: VARIANTS OF ALTERNATIVE SPLICING
; FILE REFERENCE: 2796-0160P
; CURRENT APPLICATION NUMBER: US/09/778,927A
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: IL 134453
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: IL135341
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(582)
; OTHER INFORMATION: n = a,c,g,t any unknown or other

US-09-778-927A-28

Query Match 7.2%; Score 87.2; DB 9; Length 582;
Best Local Similarity 83.1%; Pred. No. 5.7e-12;
Matches 123; Conservative 0; Mismatches 23; Indels 2; Gaps 2;

Qy	28	CAGTGGATTATATTACAACTTTTATTATTATTCAGGCCATTTCCAGCAGTTTAAAG	87
Db	436	GAATGAATGCTACAGCTATACCCAGACCTTTTATAGGCCATTTCCAGCAGTTTAAAG	495
Qy	88	GTTC-CTTTCAGCTATGAGTCTCATAGAGGCTGTAAATGTTGATACACAGTTTCA	146
Db	496	GTTCCTCTTTCAGCTATGAGTCTCATAGAGGCTGTAAATGTTGATACACAGTTTCA	554
Qy	147	ACGTCACACAGTGAAGACTTCAGAAAT	174
Db	555	ACGTCACACAGTGAAGACTTCGAAAT	582

RESULT 15

US-10-085-783A-5829
; Sequence 5829, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5829
; LENGTH: 172
; TYPE: DNA
; ORGANISM: Human

US-10-085-783A-5829

Query Match 6.2%; Score 74.8; DB 13; Length 172;
Best Local Similarity 74.7%; Pred. No. 4.5e-09;
Matches 124; Conservative 0; Mismatches 32; Indels 10; Gaps 2;

Qy	1060	CAGATCAGTTTGGGTGCAATGATGATTTTCAGCGTCAAAATGGAGTAGGGAAGATTT	1119
Db	1	CAGATCAGTTTGGGTGCAATGATGATTTTCAGCGTCAAAATGGAGTAGGGAAGATTT	60
Qy	1120	CTGTATA-----CTTGCTGGAGAGGAGGAATGTGTATAGTTACTCATTTA-----CATGA	1169
Db	61	CTGTATATACTTTGCTCGGGAGAGGAGGAATGTGTATAGAACTGATTTAGGGTGACTT	120
Qy	1170	CTCCAAAACCTTTTATTAACCAATTTTAGTTTAAACCAAAATAA	1215
Db	121	CCAAAACCTTTTGTTTTAAACCAATTTTAGTTTTCAGAAAGATAA	166

Search completed: June 16, 2004, 01:06:24
Job time : 421 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2004, 23:03:52 ; Search time 81 seconds
(without alignments)
8324.263 Million cell updates/sec

Title: US-10-009-557-34
Perfect score: 1215
Sequence: 1 gaagaactagcatgtatgta.....ttagttttaaaaaaaaaa 1215

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCRUS_COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Fred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1153.2	94.9	1471	4	US-09-620-312D-1099
2	715.8	58.9	1241	4	US-09-833-381-1401
3	56.2	4.6	2169	4	US-09-434-408-3
4	48.2	4.0	1947	4	US-09-907-794A-184
5	48.2	4.0	1947	4	US-09-905-125A-184
6	48.2	4.0	1947	4	US-09-902-775A-184
7	44.2	3.6	832	4	US-09-621-976-2813
8	44	3.6	1749	4	US-09-081-149-5
9	43.2	3.6	7218	1	US-08-232-463-14
10	42	3.5	419	4	US-09-833-381-1373
11	41.2	3.4	1881	4	US-09-833-381-1406
12	39.4	3.2	1680	4	US-09-081-149-1
13	38.4	3.2	832	4	US-09-621-976-2813
14	38.4	3.2	3401	4	US-09-907-794A-249
15	38.4	3.2	3401	4	US-09-905-125A-249
16	38.4	3.2	3401	4	US-09-902-775A-249
17	38.4	3.2	4428	4	US-09-833-381-1364
18	38.2	3.1	450	4	US-09-621-976-406
19	38.2	3.1	495	4	US-09-312-283C-361
20	38.2	3.1	5199	4	US-09-620-312D-40
21	37.8	3.1	7304	4	US-10-234-708-44
22	37.6	3.1	9370	1	US-08-322-559-27
23	37.6	3.1	9370	3	US-08-545-860D-27
24	37.6	3.1	9370	5	PCI-US94-04496-27
25	37.6	3.1	9391	1	US-08-320-559-25
26	37.6	3.1	9391	3	US-08-545-860D-25
27	37.6	3.1	9391	5	PCI-US94-04496-25

C 28	37.4	3.1	3026	4	US-09-963-137-136	Sequence 136, App
C 29	37.4	3.1	3026	4	US-09-963-137-140	Sequence 140, App
C 30	37	3.0	1749	4	US-09-081-149-6	Sequence 6, Appli
C 31	36.2	3.0	1218	4	US-09-252-991A-14561	Sequence 14561, A
C 32	36.2	3.0	5407	3	US-09-269-040-7	Sequence 7, Appli
C 33	36.2	3.0	6877	1	US-08-347-340-1	Sequence 1, Appli
C 34	36	3.0	72604	4	US-09-268-992-7	Sequence 7, Appli
C 35	36	3.0	72604	4	US-09-657-474-7	Sequence 7, Appli
C 36	36	3.0	640681	4	US-09-790-388-1	Sequence 1, Appli
C 37	35.8	2.9	260	4	US-09-833-381-557	Sequence 557, App
C 38	35.8	2.9	842	4	US-09-484-370B-16	Sequence 16, Appl
C 39	35.6	2.9	3012	4	US-09-184-418C-57	Sequence 57, Appl
C 40	35.6	2.9	8700	2	US-08-392-625-16	Sequence 16, Appl
C 41	35.6	2.9	8750	2	US-08-466-961A-16	Sequence 16, Appl
C 42	35.6	2.9	8780	2	US-08-645-193B-18	Sequence 18, Appl
C 43	35.6	2.9	8954	4	US-09-184-418C-6	Sequence 6, Appli
C 44	35.6	2.9	640681	4	US-09-790-388-1	Sequence 1, Appli
C 45	35.4	2.9	11459	4	US-09-462-136-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-620-312D-1099
; Sequence 1099, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6589662el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 1399
; LENGTH: 1471
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (40)..(1284)
US-09-620-312D-1099

Query Match 94.9%; Score 1153.2; DB 4; Length 1471;
Best Local Similarity 99.7%; Pred. NO. 1.2e-312;
Matches 1155; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 58 TTATTCAGCCATTTCCAGCAGTTTAAAGGTTTCCTTTCAGCTATGAGACTGGCTCAT 117
Db 314 TAAGTAAGCCATTTCCAGCAGTTTAAAGGTTTCCTTTCAGCTATGAGACTGGCTCAT 373

```
QY 118 GAGCTGTAATGTTGATACACCAAGTTTCAACGGCTCACACCAAGTGAAGACTTCAGAAATTG 177
Db 374 GAGCTGTAATGTTGATACACCAAGTTTCAACGGCTCACACCAAGTGAAGACTTCAGAAATTG 433
QY 178 AAAACCTTTAAACCTAAATAGTTATATCATATCAATCAACCAAAAGACTATCCTCTAAGTAAGAAAT 237
Db 434 AAAACCTTTAAACCTAAATAGTTATATCATATCAATCAACCAAAAGACTATCCTCTAAGTAAGAAAT 493
QY 238 TTCCATATTCCTTGGACATCTTCAGACTTCTTACTGTGGCTGTGGAGTTGATATGC 297
Db 494 TTCCATATTCCTTGGACATCTTCAGACTTCTTACTGTGGCTGTGGAGTTGATATGC 553
QY 298 GTATGCTTTCCTTAAAGCCTTAGGAAATAGACTTTGAGTCAACCAATATAAAAGC 357
Db 554 GTATGCTTTCCTTAAAGCCTTAGGAAATAGACTTTGAGTCAACCAATATAAAAGC 613
QY 358 TTCCAGTCAAAATGGAGACTCATACACCTTCAAGAACTTAACTGTAATGACAACTCACT 417
Db 614 TTCCAGTCAAAATGGAGACTCATACACCTTCAAGAACTTAACTGTAATGACAACTCACT 673
QY 418 TGGAGTCAATTTAGTGTAGCTTGTGTCATCTACCTCCAGAGTCACTTGGAGTTGG 477
Db 674 TGGAGTCAATTTAGTGTAGCTTGTGTCATCTACCTCCAGAGTCACTTGGAGTTGG 733
QY 478 ACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGCAAGTTTGGCAGCTCCAGGAACCTTA 537
Db 734 ACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGCAAGTTTGGCAGCTCCAGGAACCTTA 793
QY 538 AGAATTTAAACCTTGAGTAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 597
Db 794 AGAATTTAAACCTTGAGTAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 853
QY 598 TAAACCTTCGCTTTTGTGACGAGCTCGAATAGCTTCCATTTTCCCTAGTGAATTTA 657
Db 854 TAAACCTTCGCTTTTGTGACGAGCTCGAATAGCTTCCATTTTCCCTAGTGAATTTA 913
QY 658 GAAATTTATCCCTTGAATATCTGATCTTTTGGAAATATCTTTTGAACCAACCAAGTCC 717
Db 914 GAAATTTATCCCTTGAATATCTGATCTTTTGGAAATATCTTTTGAACCAACCAAGTCC 973
QY 718 TTCCAGTAATAAGCTGCAAGCACTTAATCTTATTTGGAATCTTCTGCGAAGCAACATAT 777
Db 974 TTCCAGTAATAAGCTGCAAGCACTTAATCTTATTTGGAATCTTCTGCGAAGCAACATAT 1033
QY 778 TACATAATAGGATTCATATGGCTCTCATATCATTCATTCATCTCTGCAAGATTTGG 837
Db 1034 TACATAATAGGATTCATATGGCTCTCATATCATTCATTCATCTCTGCAAGATTTGG 1093
QY 838 ATACCGCAAAATTTGTTGTTGGAAGATTTCTGCTGAACTCTTTCAATTCAGGAACCTA 897
Db 1094 ATACCGCAAAATTTGTTGTTGGAAGATTTCTGCTGAACTCTTTCAATTCAGGAACCTA 1153
QY 898 CTACCATGAATCTGCAATCTGTGTCACCACTGCTGTCTTAGTAGATTAATTTGGTGGTA 957
Db 1154 CTACCATGAATCTGCAATCTGTGTCACCACTGCTGTCTTAGTAGATTAATTTGGTGGTA 1213
QY 958 CTGAGCACTATATCTCTTATTTCTGTTCTCTAGCTGTGTATGTTAATTTCTCTGATA 1017
Db 1214 CTGAGCACTATATCTCTTATTTCTGTTCTCTAGCTGTGTATGTTAATTTCTCTGATA 1273
QY 1018 TGTAAAGTAATGGGTGAGACCAAGAAAAGAAATTTCAATAACAGATCAGTTTGGGGTGC 1077
Db 1274 TGTAAAGTAATGGGTGAGACCAAGAAAAGAAATTTCAATAACAGATCAGTTTGGGGTGC 1333
QY 1078 ATGTATGATTTTGCAGGCTCAATTTGGAGTAAGGGAAGATTTCTGTATCTTCTGAGTA 1137
Db 1334 ATGTATGATTTTGCAGGCTCAATTTGGAGTAAGGGAAGATTTCTGTATCTTCTGAGTA 1393
QY 1138 GGAGGAATGTATAGTTACTCATTTAGATGATCCAAACCTTTTATTAATAACCAATTTT 1197
Db 1394 GGAGGAATGTATAGTTACTCATTTAGATGATCCAAACCTTTTATTAATAACCAATTTT 1453
QY 1198 AGTTTAAAAAATAAAAA 1215
```

```
Db 1454 AGTTTAAAAAATAAAAA 1471
RESULT 2
US-09-833-381-1401
; Sequence 1401, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1401
; LENGTH: 1241
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1401
Query Match 58.9%; Score 715.8; DB 4; Length 1241;
Best Local Similarity 97.6%; Pred. No. 2e-190;
Matches 737; Conservative 0; Mismatches 17; Indels 1; Gaps 1;
QY 58 TTATTCAGGCGCAATTCACGAGCTTTAAAAAGGTTTCCCTTCAGCTATGAGACTGGCTCATTA 117
Db 371 TAAGTAAGGCGCAATTCACGAGCTTTAAAAAGGTTTCCCTTCAGCTATGAGACTGGCTCATTA 430
QY 118 GAGCTGTAATGTTGATACACCAAGTTTCAACGGCTCACACCAAGTGAAGACTTCAGAAATTG 177
Db 431 GAGCTGTAATGTTGATACACCAAGTTTCAACGGCTCACACCAAGTGAAGACTTCAGAAATTG 490
QY 178 AAAACCTTTAAACCTAAATAGTTATATCATATCAATCAACCAAAAGACTATCCTCTAAGTAAGAAAT 237
Db 491 AAAACCTTTAAACCTAAATAGTTATATCATATCAATCAACCAAAAGACTATCCTCTAAGTAAGAAAT 550
QY 238 TTCCATATTCCTTGGAACTCTTCAGACTTCTTACTGTGGGCTTGTCCGAGTTGATATGC 297
Db 551 TTCCATATTCCTTGGAACTCTTCAGACTTCTTACTGTGGGCTTGTCCGAGTTGATATGC 610
QY 298 GTATGCTTTCCTTAAAGCCTTAGGAAATAGACTTTGAGTCAACCAATATAAAAGC 357
Db 611 GTATGCTTTCCTTAAAGCCTTAGGAAATAGACTTTGAGTCAACCAATATAAAAGC 670
QY 358 TTCCAGTCAAAATGGAGACTCATACACCTTCAAGAACTTAAACCTGAAATGACAACTCACT 417
Db 671 TTCCAGTCAAAATGGAGACTCATACACCTTCAAGAACTTAAACCTGAAATGACAACTCACT 730
QY 418 TGGAGTCAATTTAGTGTAGCTTGTGTCATTCATCTCAGTCCAGAAAGTCACTTCGAGTTGG 477
Db 731 TGGAGTCAATTTAGTGTAGCTTGTGTCATTCATCTCAGTCCAGAAAGTCACTTCGAGTTGG 790
QY 478 ACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGCAAGTTTGGCAGCTCCAGGAACCTTA 537
Db 791 ACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGCAAGTTTGGCAGCTCCAGGAACCTTA 850
QY 538 AGAATTTAAACCTTGA- CGATAATGAATTTGATTCATTTCTTGGAGATAGGCAACTTA 596
Db 851 AGAATTTAAACCTTGA- CGATAATGAATTTGATTCATTTCTTGGAGATAGGCAACTTA 910
QY 597 ATAAACCTTCGCTTTTGTGACGAGCTCGAAATTAAGCTTTCCATTTTGTGCTAGTGAATTT 656
Db 911 ATAAACCTTCGCTTTTGTGACGAGCTCGAAATTAAGCTTTCCATTTTGTGCTAGTGAATTT 970
QY 657 AGAATTTATCCCTTCAATCTTGGATCTTTTGGAAATACCTTTTGAACCAACCAAGTCC 716
Db 971 AGAATTTATCCCTTCAATCTTGGATCTTTTGGAAATACCTTTTGAACCAACCAAGTCC 1030
QY 717 CTTCAGTAATAAAGCTGCAAGCAACCACTTAACCTTTATTTGATCTTCTTCGACCAACCTTA 776
```

Db 1031 CTTCCAGTAATAAGCTGCAAGCACCATTAACTTTATTGGAATCTTCTGCAAGAACCATTA 1090
Qy 777 TTACATATAGATGTCATATATGCTTCATATCAT 811
Db 1091 TTACATAATAGTAAGATATTTAATAGTCATGTAAT 1125

RESULT 3

US-09-434-408-3
; Sequence 3, Application US/09434408
; Patent No. 6440697
; GENERAL INFORMATION:
; APPLICANT: Venezia, Domenick
; APPLICANT: Grossmann, Angelika
; TITLE OF INVENTION: RING FINGER PROTEIN ZAPOP3
; FILE REFERENCE: 98-41
; CURRENT APPLICATION NUMBER: US/09/434,408
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: US 60/108,258
; EARLIER FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2169
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate polynucleotide sequence of zapop3
; NAME/KEY: misc feature
; LOCATION: (1)...(2169)
; OTHER INFORMATION: n = A,T,C or G
US-09-434-408-3

Query Match 4.6%; Score 56.2; DB 4; Length 2169;
Best Local Similarity 30.4%; Pred. No. 6.5e-06;
Matches 109; Conservative 57; Mismatches 183; Indels 9; Gaps 1;

Qy 301 TCGTTGCTTAAAGCCCTTAGGAATAGCTTGAGTCACACACCATATATAAAGCTTC 360
Db 230 TTTTWSNTYNGCNACNATHAAGTNTYNGAYTNCAYGAYTNCAYTNCNGCNTNC 289
Qy 361 CAGCTACATATGAGACCTCATACACCTTCAAGACCTTAACCTGAATGATCACTTGG 420
Db 290 CNGAYGAYTNGCNCARYTNCNGCNTNCARGTNTYNGAYTNGARMGNAAYCARYTNA 349
Qy 421 AGTCATTTAGTGTAGCTTGTGTCATCTACACTCAGAGTCACCTGGAGTTGGACC 480
Db 350 TGCARYTNCNMGNSNATHGGNAAYTNCNCA-----RYTNCARACNTYTNAYG 400
Qy 481 TCAGCAAGAACAAATCAAGGCACCTCCCTGTGCGAGTTTTCAGCTCCAGGAACCTTAAGA 540
Db 401 TNAAGAYAAAYATYNAARGARYTNCNGAYCNGCTNGNGARYTNGWSNYTNMGNA 460
Qy 541 APTTAAACTTCAGCATATAGTAATGATTCATTTCTTCAGATAGGACAACTAATAA 600
Db 461 CNYTNAAYATHSNGGNAAYGARATHCARMGNTNCNCARATGYTNGCNCAYTNMGNA 520
Qy 601 ACCTTCGCTTTTGTGCAGACCTCGAAATAGCTTCATTTTGCCTAGTGAATTAG 658
Db 521 CNYTNGARATGYTNWSNTYNGAYGNCNCGNATGTTNTAYCCNCGMNGARGTNTG 578

RESULT 4

US-09-907-794A-184
; Sequence 184, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Inc

; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Pilvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 184
; LENGTH: 1947
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-907-794A-184

Query Match 4.0%; Score 48.2; DB 4; Length 1947;
Best Local Similarity 48.7%; Pred. No. 0.0011;
Matches 168; Conservative 0; Mismatches 168; Indels 9; Gaps 1;

Qy 248 CTTGGAACTCTTCAGACTTCTTACTGTGGCTGTCCGAGTTCGATATCGTATGCTTTG 307

Db 1121 CTGGAGTCACTTATTTCTCTAACCAAGCTCGAATCTTACAGTGGCAGTATTAG 1180
Qy 308 CTTAAAGCCTTAGGAATAGACTTGGAGTCAAGCAATATAAAGAGCTTCCAGCTAC 367
Db 1181 TTTACAGAACTCGATGCTTAGATGTGAGCTCAACAACATTTCAATGATTTCAATAGA 1240
Qy 368 AATTGGAGACCTCATACACCTTCAAGAACTTAACTGAATGCAATCACTTTGGAGTCAAT 427
Db 1241 AATAGGATTGCTTCCAGAACCTTGAGCAATTTGCATATCACTGGGAACAAGTGGACATTTCT 1300
Qy 428 TAGTGTAGCCTTGTGTGATCTTCACTCCAGAGTCACTTCCGAGTGGACCTCAGCAA 487
Db 1301 GCCAAAACAAATTTTAAATGCATAAAGTTGA-----GGACTTTGAAATCTGGGACA 1351
Qy 488 GAACAAATCAAGCACTCCCTGTGGAGTTTGGCAGCTCCAGGAACTTAAGAAATTTAAA 547
Db 1352 GAATGATCACTCACTCCAGAGAAAGTTGGTCACTTCCAGCTCACTCAGCTGGA 1411
Qy 548 ACTTGAGGATAATGAATGATTCAATTTCCCTTGCAGATAGGACA 592
Db 1412 GCTGAAGGGGAACTGCTTGGACCGCTGCCAGCCAGCTGGGCA 1456

RESULT 5

US-09-905-125A-184
; Sequence 184, Application US/09905125A

; Patent No. 6664376

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Goddard, A.

; APPLICANT: Gottsen, Mary E.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth, J.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Mather, Jennie P.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/905,125A

; CURRENT FILING DATE: 2001-07-12

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547

; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 184
; LENGTH: 1947
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-905-125A-184

Query Match 4.0%; Score 48.2; DB 4; Length 1947;
Best Local Similarity 48.7%; Pred. No. 0.0011;
Matches 168; Conservative 0; Mismatches 168; Indels 9; Gaps 1;
Qy 248 CTTGGAACATCTTCAGACTTCTTACTGTGGCTTGTCCGAGTTGATATGCGTATGCTTTG 307
Db 1121 CTTGAGTCACTTATTTCTCTAACCAAGCTCGAATCTTACAGTGGCAGTATTAG 1180
Qy 308 CTTAAAGCCTTAGGAATAGACTTGGAGTCAAGCAATATAAAGAGCTTCCAGCTAC 367
Db 1181 TTTACAGAACTCGATGCTTAGATGTGAGCTCAACAACATTTCAATGATTTCAATAGA 1240
Qy 368 AATTGGAGACCTCATACACCTTCAAGAACTTAACTGAATGCAATCACTTTGGAGTCAAT 427
Db 1241 AATAGGATTGCTTCCAGAACCTTGAGCAATTTGCATATCACTGGGAACAAGTGGACATTTCT 1300
Qy 428 TAGTGTAGCCTTGTGTGATCTTCACTCCAGAGTCACTTCCGAGTGGACCTCAGCAA 487
Db 1301 GCCAAAACAAATTTTAAATGCATAAAGTTGA-----GGACTTTGAAATCTGGGACA 1351
Qy 488 GAACAAATCAAGCACTCCCTGTGGAGTTTGGCAGCTCCAGGAACTTAAGAAATTTAAA 547
Db 1352 GAATGATCACTCACTCCAGAGAAAGTTGGTCACTTCCAGCTCACTCAGCTGGA 1411
Qy 548 ACTTGAGGATAATGAATGATTCAATTTCCCTTGCAGATAGGACA 592
Db 1412 GCTGAAGGGGAACTGCTTGGACCGCTGCCAGCCAGCTGGGCA 1456

RESULT 6

US-09-902-775A-184

; Sequence 184, Application US/09902775A

; Patent No. 6686451

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gottsen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/902,775A
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00215
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 184
LENGTH: 1947
TYPE: DNA
ORGANISM: Homo sapiens
US-09-902-775A-184

Query Match 4.0%; Score 48.2; DB 4; Length 1947;
Best Local Similarity 48.7%; Pred. No. 0.0011;
Matches 168; Conservative 0; Mismatches 168; Indels 9; Gaps 1;

QY 248 CTGGAACATCTTCAGACTCTTACTGTGGGCTTGTCGAGTTGATATGCGTATGCTTTG 307
DB 1121 CTGAGGTCATCTTTATTTCTCTAACACAGCTCGAATCTTACAGTGGCAGTATTTAG 1180

QY 308 CTTAAAAAGCCTTAGGAATTAGACTTAGTCAACCAATATAAAAAAGCTTCCAGCTAC 367
DB 1181 TTTACAGAACTCAGATGCTTAGATGTGAGCTACAAACAATTTCAATGATTTCCAAATAGA 1240

QY 368 AATTGGAGACCTCATACCTTTCAGAACTTAACTGAATCAATCACTTGGAGTCAAT 427
DB 1241 AATAGGATTGCTTCAGAACCTGCAGCATTTTGATATCACTCGGGAACAAAGTGCATCT 1300

QY 428 TAGTGTAGCCTTGTCTCATTCTACACTCCAGAGTCACTTCGAGTTTGAGCCTCAGCA 487
DB 1301 GCCAAACAATTTGTTTAAATGCATAAAGTTGA-----GCACTTTGAATCTGGACA 1351
QY 488 GAACAAAATCAAGGCACCTCCCTGTGCAGTTTTCAGAGTCCAGGAACCTTAAGAATTTAA 547
DB 1352 GAACCTGATCCTCCTCCTCCAGAGAAAGTTGTCAGCTCTCCAGCTCCTCAGCTGGA 1411
QY 548 ACTGACGATATGAATTAATTCATTCCTTCCTGGAAGATAGACA 592
DB 1412 GCTGAAGGGGAACCTGCTGGACCGCTGCCAGCCAGCTGGGCCA 1456

RESULT 7

US-09-621-976-2813
; Sequence 2813, Application: US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match 3.6%; Score 44.2; DB 4; Length 832;

Best Local Similarity 10.8%; Pred. No. 0.0094;
Matches 39; Conservative 176; Mismatches 139; Indels 7; Gaps 1;

QY 854 TGTTGTGGAAGATTCTGTCTGAATCTTTCAATCAAGGAATCTACCATGAATCTGCA 913
DB 17 TKWKSWSYMYKWKYMYKTYWRNRKXKKKAWKWKYKWTWYRYMYWYKWKYKWKAMCET 76
QY 914 TTCTGTGCCCCACACTGTGCTTAGTAGATAATTGGGTGCTACTGAGACCTATTAT 973
DB 77 KTKKKKKGYMMWYWGWRSSYMAWTRCTGYAYRSMYWRKWKWKYKAYIRKITCY 136
QY 974 CTCCTATTCTCTCTAGGCTGTATCTTAATCTCTGATATCTTAAAGTAATGSGT 1033
DB 137 SSKGWTWKKKKKANTTWKKYKYYAATRYMMWMTKWRASWYCWYKWKARWSTM 196
QY 1034 GAGACCAAGAAAGAAATTTCAATACAGATCAGTTTGGGTGCTATGATTTTGGAG 1093
DB 197 RKRSYSARSARCC-----YSCWGAWSWKYMYWRWRWATGAGMKAWRASCMW 249
QY 1094 COTCAANTTGGAGTAGGGAAGATTCTGTATATCTTCTGAGAGAGGAGGAATGTGTATAG 1153
DB 250 RKKYAGKSTSYKSMWTRSWKTYCTKARVIGYCYKGGMWGRGRWYKWKYKWKAR 309
QY 1154 TPTACTCATTTAGATGACTCCAAACCTTTTATTAACCAATTTTAGTTTAAAAA 1213
DB 310 WWWCWARMYRSTGTGRASWNRWYTYTMMKWKYAWAFARWMMWMAWERRACAAAT 369
QY 1214 A 1214
DB 370 A 370

RESULT 8

US-09-081-149-5
; Sequence 5, Application US/0908149A

```

? FILING DATE:
? APPLICATION NUMBER: EP 91 114 300.6
? FILING DATE: 26-AUG-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: BENT, Stephen A.
? REGISTRATION NUMBER: 29,768
? REFERENCE/DOCKET NUMBER: 30472/114 IMMU
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (703)836-9300
? TELEFAX: (703)683-4109
? TELEX: 899149
? INFORMATION FOR SEQ ID NO: 14:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 7218 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? IMMEDIATE SOURCE:
? CLONE: pTZgpt-Fls
?
? US-08-232-463-14
?
? Query Match 3.6%; Score 43.2; DB 1; Length 7218;
? Best Local Similarity 8.1%; Pred. No. 0.049;
? Matches 36; Conservative 210; Mismatches 198; Indels 0; Gaps 0;
?
? QY 601 ACCTCGCTTTTGTGACGAGCTCGAAATAAGCTTCCATTTTGCTAGTAGAATTAGAA 660
? DB 1036 AGCTTGGCTGCAGCTCGAGGAGCTTGCATTTTTTTTTTTTTTTTTTTTTTTTTT 1095
? QY 661 ATTATCCCTTGAAATCTTTGGATCTTTTGGAAATCTTTTGAACAACCAAGTCTCTC 720
? DB 1096 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1155
? QY 721 CAGTAATAAGCTGCAAGCACACATTAACCTTTATGGAATCTCTCGACGACCAATATAC 780
? DB 1156 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1215
? QY 781 ATAATAGGATTCATATGCTCTCATATCATTCATTCATTCATTCGCAAGATTGGATA 840
? DB 1216 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1275
? QY 841 CGCAAAATTTGTTGTTGGAGATCTGCTGMACTCTCTTCACTCAAGGAACCTACTA 900
? DB 1276 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1335
? QY 901 CCATGAATCGCATCTGTTGCCACACTGCTCTTAGTAGATAATTTGGGTGGTACTG 960
? DB 1336 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1395
? QY 961 AAGCACTAATCTCTAATTTCTGTTCTCTAGGCTGTATGTTAATCTCTCTGATATG 1020
? DB 1396 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYGGTACCAATCTCTATCTCT 1455
?
? QY 1021 TAAAGTAATCGGTGACACAGAA 1044
? DB 1456 TTAACCTACTTCATAGTAGGTA 1479
?
? RESULT 10
? US-09-833-381-1373
? Sequence 1373, Application US/09833381
? Patent No. 6672186
? GENERAL INFORMATION:
? APPLICANT: Robison, Keith E.
? TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
? FILE REFERENCE: 5800-119
? CURRENT APPLICATION NUMBER: US/09/833,381
? PRIOR FILING DATE: 2001-04-11
? PRIOR APPLICATION NUMBER: 09/516,448
? PRIOR FILING DATE: 2000-02-29
? NUMBER OF SEQ ID NOS: 205C
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 1373

```



```
; LENGTH: 419
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1373

Query Match      3.5%; Score 42; DB 4; Length 419;
Best Local Similarity 60.5%; Pred. No. 0.028;
Matches 69; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 314 AGCCCTTAGAATTAGACTTGAGTGCACACCACTATATAAAAGCTTCACGCTCAATGG 373
Db 246 AGAGCTAGAGACACCTTGATTTAGCAACCAATGAATATATATATATTTGCGCAATCAATGG 305
QY 374 AGACCTCATACACCTTCAAGACTTAACTGGAATCAATCAATCACTTGGAGTCATT 427
Db 306 AGCCCTTACATCTAATAAGATCTCTGGTGGATGGAAATCAATCACTGTCAGAAATT 359

RESULT 11
US-09-833-381-1406
; Sequence 1406, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith B.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1406
; LENGTH: 1881
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1406

Query Match      3.4%; Score 41.2; DB 4; Length 1881;
Best Local Similarity 53.8%; Pred. No. 0.095;
Matches 85; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 305 TTGCTTAAAGCCCTTAGAATTAAGACTTGATGATCAACCACTATATAAAAGCTTCACG 364
Db 339 TGGCCTCACACACCTGGTGAAGCTAGACCTGAGTAGAAGAACAGCTGCACAGCTGCCAGC 458
QY 365 TACAATCGAGACCTCATACACCTTCAAGAACTTAACCTGAATGACATCACTTGGAGTC 424
Db 459 AGACTTGGCGTCTGGTCAACCTCCAGCACCCTGGATCTCCTCAACCAACAGCTGTCAC 518
QY 425 ATTAGTGTAGCTTGTGTCTATTCACACTCCAGAACT 462
Db 519 CTGCTGTGACGCTTGTCTACGCTCAAGAACCTGAACT 556

RESULT 12
US-09-081-149-1
; Sequence 1, Application US/09081149A
; Patent No. 6506889
; GENERAL INFORMATION:
; APPLICANT: Han, Min
; APPLICANT: Sieburth, Derek
; TITLE OF INVENTION: RAS SUPPRESSOR SUR-8
; FILE REFERENCE: UTC-02938
; CURRENT APPLICATION NUMBER: US/09/081,149A
; CURRENT FILING DATE: 1998-05-19
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1680
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-081-149-1

Query Match      3.2%; Score 39.4; DB 4; Length 1680;
Best Local Similarity 65.2%; Pred. No. 0.29;
Matches 58; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 340 ACAACCATATAAAAAAGCTTCCAGCTACAATTTGGAGAGCTCATACACCTTCAAGAACTTA 399
Db 1220 ACATCACTCACTGAAAAGCTTCCAAATCAATAGGAATCTCATATAAACTCCGAGCTGG 1279
QY 400 ACTGGAATGACAATCACTTGGAGTCATTT 428
Db 1280 ATCTCGAGGAATGAATGGAGACCTT 1308

RESULT 13
US-09-621-976-2813/c
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match      3.2%; Score 38.4; DB 4; Length 832;
Best Local Similarity 10.4%; Pred. No. 0.39;
Matches 36; Conservative 157; Mismatches 153; Indels 0; Gaps 0;

QY 850 TTGTGTTTGTGGAAGATCTGTCTGAACCTCTTCACTCAAGCACTTACCATCAATC 909
Db 347 TWTMMWKKKARRYYWKKSTVACASRYKRTWGWYWWKEMMSTRWYCYWCKCCMY 288
QY 910 TGCATCTCTTGGCCACACTGGCTCTTAGTAGATAATTTGGGTGGTACTGAAGACACCTA 969
Db 287 RGERCAWYTWARGRWMSYAWGKWSKMSRSMCTRYKKGSTYWTMKCTCATWYWK 228
QY 970 TTATCTCTTATTTCTGTTCTCTAGGCTGTTATGTTTAAATTCCTCTGATATGTTAAAGTAAT 1029
Db 227 YKRWMSKTCWSGRGGYNTSYTSTSYMYWASWYTWGWWGRWWSYTWYVWKKW 166
QY 1030 GGTGAGACCAAGAAATTTCAATAACAGATCAGTTTGGGTGTCATGATGATTTT 1089
Db 167 RYATTWRRAMWMAAATWMMYMWNAWCKSSRGAAMYRRTTMMWGYRWKKSYYRTRC 108
QY 1090 GCAGCGTCAAACTGGAGTAAGGAAGATTTCTGTATATCTCTGGAGAGAGAAATGTGT 1149
Db 107 AWAYAKTKRSYWCWKKKCCMMMMMAAAYGKTMNRACWKTRYWMAWMAWMAW 48
QY 1150 ATAGTTACTCATTTAGATGACTCCAAAACCTTTTATTAACCAAT 1195
Db 47 TMMWYWWRAMKRWKRWKRSWWSWMAWGYTRWAAWMAWRY 2

RESULT 14
US-09-794A-249
; Sequence 249, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
```

APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kijavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/907,794A
CURRENT FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 249
LENGTH: 3401
TYPE: DNA
ORGANISM: Homo Sapien
US-09-907-794A-249

Query Match 3.2%; Score 38.4; DB 4; Length 3401;
Best Local Similarity 53.3%; Pred. No. 0.75;
Matches 81; Conservative 0; Mismatches 71; Indels 3; Gaps 0;

QY 453 CTCGAGAGTCACTTCGAGTTTGGACCTCAGCAGACAAAATCAAGGCTCCCTGTG 512
Db 1403 CTCTGCAGAACCTCCAGAACCTAGCCATCAGGCCAACCGGATCGAGACGCTCCCTCG 1462
QY 513 CAGTTTTCGCCAGCTCCAGGAACCTTAAGAATTTAAAACTTGACGATATGAATTTGATTCAA 572
Db 1463 GAGCTCTTCGAGTGCAGGAAGCTGCGGCGCTGCACCTGGGCACACAGCTGCTGCAGTCA 1522
QY 573 TTTCCTTGCAAGATAGGACAACTAATAAACCT 604
Db 1523 CTGCCCTCCAGGTGGCGAGCTGACCAACCT 1554

RESULT 15

US-09-905-125A-249
; Sequence 249, Application US/09905125A
; Patent No. 6664376
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/905,125A
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565

```
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 249
; LENGTH: 3401
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-905-125A-249

Query Match      3.2%; Score 38.4; DB 4; Length 3401;
Best Local Similarity 53.3%; Pred. No. 0.75;
Matches 81; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY      453 CTCAGAAGTCACCTCGGAGCTTTGGACCTCAGCAAGAACAAATCAAGGCACTCCCTGTG 512
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1403 CTCCTGCAGAACCTCCAGAACCTAGCCATCAGGCCAACCGGATCGAGCGCTCCCTCG 1462
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      513 CAGTTTGGCCAGCTCCAGGAACCTTAAGAAATTTAAACTTCAGCGATATGAATTGANTCAA 572
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1463 GAGCTCTTCAGTGCCGGAAGCTGCGGGCCCTGCACCTGGGCAACACGCTGCTGCAGTCA 1522
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      573 TTTCCTTGCAGATAGACAACTAATAAACCT 604
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1523 CTGCCCTCCAGGGTGGCGGAGCTGACCAACCT 1554
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: June 16, 2004, 00:59:05
Job time : 83 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2004, 19:53:27 ; Search time 399 Seconds
(without alignments)
12936.242 Million cell updates/sec

Title: US-10-009-557-34
Perfect score: 1215
Sequence: 1 gaagaactagatgatgtgta.....ttagttaaaaaaaaaaaaaa 1215

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04: *
1: geneseqm:980s: *
2: geneseqm:990s: *
3: geneseqm:2000s: *
4: geneseqm:2001as: *
5: geneseqm:2001bs: *
6: geneseqm:2002s: *
7: geneseqm:2003as: *
8: geneseqm:2003bs: *
9: geneseqm:2003cs: *
10: geneseqm:2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1215	100.0	1215	4 AAC66898	Aac66898 Human EXM
2	1153.2	94.9	1471	4 AAI59205	Aai59205 Human pol
3	1153.2	94.9	1471	8 Adb49189	Adb49189 Novel hum
4	1153.2	94.9	1600	6 ABK90911	Abk90911 CDNA enco
5	1139	93.7	1341	4 AAS33220	Aas33220 DNA enco
6	951.4	78.3	1287	4 AAI60991	Aai60991 Human pol
7	950.2	78.2	1251	6 AAD27818	Aad27818 Human leu
8	777	64.0	780	6 AAD27817	Aad27817 Human leu
9	521.6	42.9	682	4 AAS56409	Aas56409 Human CDN
10	443.8	36.5	467	6 ABZ68766	Abz68766 Kidney ca
11	425.2	35.0	1234	7 ABZ36181	Abz36181 Human ca
12	418.8	34.5	584	4 AAH09329	Aah09329 Human CDN
13	418.8	34.5	1590	4 AAH15353	Aah15353 Human CDN
14	331	27.2	333	7 ABZ19104	Abz19104 Group III
15	281.4	23.2	750	2 AAX98737	Aax98737 Human val
16	281	23.1	732	2 AAX98741	Aax98741 Human val
17	188.8	15.5	770	2 AAZ16653	Aaz16653 Human gen
18	91.6	7.5	430	5 AAF64752	Aaf64752 Novel hum
19	87.2	7.2	582	6 ABSS1842	Abss1842 Novel hum
20	85.2	7.0	382	5 AAF64860	Aaf64860 Novel hum
21	60.2	5.0	2950	4 AAS31006	Aas31006 Human dia
22	59.6	4.9	2156	7 ACD13337	Acd13337 Human DNA
23	56.2	4.6	2169	3 AAA30203	Aaa30203 Human RIN

24	56.2	4.6	2169	6 ABS54163	Abs54163 Human deg
25	55	4.5	4984	6 AAL49657	Aal49657 Human neu
26	54.4	4.5	2694	3 AAZ45671	Aaz45671 Nucleotid
27	53.4	4.4	1867	6 ABQ99593	Abq99593 Human cod
28	52.4	4.3	2620	4 AAH14529	Aah14529 Human CDK
29	50.8	4.2	2000	7 ADA71938	Ada71938 Rice gene
30	50	4.1	50	6 ABZ07590	Abz07590 Human leu
31	49.8	4.1	2410	4 AAH14179	Aah14179 Human CDK
32	49.8	4.1	2410	6 AAL49327	Aal49327 Human sig
33	48.2	4.0	1593	4 AAS33111	Aas33111 DNA enco
34	48.2	4.0	1630	4 AAI61075	Aai61075 Human pol
35	48.2	4.0	1947	2 AAX52247	Aax52247 Protein P
36	48.2	4.0	1947	3 ADC78504	Adc78504 Human PRO
37	48.2	4.0	1947	4 AAF72405	Aaf72405 Human PRO
38	48.2	4.0	1947	7 ACA59047	Aca59047 Human PRO
39	48.2	4.0	1947	7 ACA58444	Aca58444 CDNA enco
40	48.2	4.0	1947	7 ACA60151	Aca60151 Human CDN
41	48.2	4.0	1947	7 ACD07551	Acd07551 Novel tum
42	48.2	4.0	1947	7 ABX71599	Abx71599 Human CDN
43	48.2	4.0	1947	7 ACH06931	Ach06931 Human sec
44	48.2	4.0	1947	7 ABX96188	Abx96188 Human sec
45	48.2	4.0	1947	7 ACA05489	Aca05489 CDNA enco

ALIGNMENTS

RESULT 1
AAC66898
ID AAC66898 standard; cDNA; 1215 BP.
XX
AC AAC66898;
XX
DT 27-MAR-2001 (first entry)
XX
DE Human EXMAD-9 coding sequence SEQ ID NO: 34.
XX
XX
XX Extracellular matrix and adhesion-associated protein; EXMAD; cancer;
XX Inflammation; reproductive disorder; cardiovascular disorder;
XX Immune disorder; musculoskeletal disorder; developmental disorder;
XX Gastrointestinal disorder; cell proliferation disorder; ss.
XX Homo sapiens.
XX
XX WO200068380-A2.
XX
PD 16-NOV-2000.
XX
XX 10-MAY-2000; 2000WO-US012811.
XX
PR 11-MAY-1999; 99US-0133643P.
PR 23-AUG-1999; 99US-0150409P.
XX
(INCY-) INCYTE GENOMICS INC.
XX Bandman O, Hillman JL, Tang YT, Lal P, Yue H, Baughn MR, Lu DM;
PI Azimzai Y;
XX
XX WPI; 2001-007395/01.
DR P-PSDB; AAB27231.
XX
PT Isolated polynucleotide encoding extracellular matrix or adhesion-associated protein (EXMAD) useful for diagnosing, treating, or preventing disorders associated with expression of EXMAD such as proliferative, immune and genetic disorders.
PT
PS Claim 4; Page 118; 129pp; English.
XX
CC The present invention provides the protein and coding sequences for 25 novel extracellular matrix and adhesion-associated proteins (EXMADs).
CC These are designated EXMAD-1, EXMAD-2, EXMAD-3, EXMAD-4, EXMAD-5, EXMAD-6, EXMAD-7, EXMAD-8, EXMAD-9, EXMAD-10, EXMAD-11, EXMAD-12, EXMAD-13, EXMAD-14, EXMAD-15, EXMAD-16, EXMAD-17, EXMAD-18, EXMAD-19, EXMAD-20, EXMAD-21, EXMAD-22, EXMAD-23, EXMAD-24, EXMAD-25.

CC EXMAD-21, EXMAD-22, EXMAD-23, EXMAD-24 and EXMAD-25. They are useful in
 CC the prevention and treatment of cancers, cell proliferation,
 CC cardiovascular, reproductive, immune, musculoskeletal, developmental and
 CC gastrointestinal disorders and inflammation
 XX
 SQ Sequence 1215 BP; 381 A; 231 C; 210 G; 393 T; 0 U; 0 Other;

Query Match 100.0%; Score 1215; DB 4; Length 1215;
 Best Local Similarity 100.0%; Pred. No. 2.3e-306;
 Matches 1215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAGAACTAGCATGTATATATCTCCAGTGGAAATTAATATCTCAACCTTTTATTA 60
 DB |||||
 QY 1 GAAGAACTAGCATGTATATATCTCCAGTGGAAATTAATATCTCAACCTTTTATTA 60
 DB |||||

QY 61 TTCAGGCCATTTCCAGCAGTTTAAAGGTTTCCCTTCAGCTATGAGACTGCTCATAGAG 120
 DB |||||

QY 61 TTCAGGCCATTTCCAGCAGTTTAAAGGTTTCCCTTCAGCTATGAGACTGCTCATAGAG 120
 DB |||||

QY 121 GCTGTAATGTGTATACACCAAGTTTCAACGCTCACACAGTGAAGACTTCAGAAATTCGAA 180
 DB |||||

QY 121 GCTGTAATGTGTATACACCAAGTTTCAACGCTCACACAGTGAAGACTTCAGAAATTCGAA 180
 DB |||||

QY 181 ACTTTAAACTAAATAGTTTATCACATCCAAAAGACTATCCTCTAAGTAAGAAATTTTC 240
 DB |||||

QY 181 ACTTTAAACTAAATAGTTTATCACATCCAAAAGACTATCCTCTAAGTAAGAAATTTTC 240
 DB |||||

QY 241 CATATTCCTTGGAAACATCTTCAGACTTCTTACTGTGGGCTTGCCGAGTTGATGCGTA 300
 DB |||||

QY 241 CATATTCCTTGGAAACATCTTCAGACTTCTTACTGTGGGCTTGCCGAGTTGATGCGTA 300
 DB |||||

QY 301 TGCCTTTGCTTAAAGGCTTTAGGAATTTAGACTTGAGTCAACCATATATAAAGCTTC 360
 DB |||||

QY 301 TGCCTTTGCTTAAAGGCTTTAGGAATTTAGACTTGAGTCAACCATATATAAAGCTTC 360
 DB |||||

QY 361 CAGCTACAATTTGAGAGCTTCATACACCTTCAAGAACTTAACCTGAATGACAACTACCTGG 420
 DB |||||

QY 361 CAGCTACAATTTGAGAGCTTCATACACCTTCAAGAACTTAACCTGAATGACAACTACCTGG 420
 DB |||||

QY 421 AGTCATTTAGTGTAGCTTGTGTGTCATCTACCTCAGAACTCACTTCGGAGTTGGACC 480
 DB |||||

QY 421 AGTCATTTAGTGTAGCTTGTGTGTCATCTACCTCAGAACTCACTTCGGAGTTGGACC 480
 DB |||||

QY 481 TCAGCAAGCAACAAATCAAGGCACTCCCTGTGAGTTTCCAGCTCCAGAACTCAAGA 540
 DB |||||

QY 481 TCAGCAAGCAACAAATCAAGGCACTCCCTGTGAGTTTCCAGCTCCAGAACTCAAGA 540
 DB |||||

QY 541 ATTATAAACTTGACGATATGAATGAATTCATTTCTTGCAGATAGGACAACTAATAA 600
 DB |||||

QY 541 ATTATAAACTTGACGATATGAATGAATTCATTTCTTGCAGATAGGACAACTAATAA 600
 DB |||||

QY 601 ACCTTCGCTTTTGTGTCAGAGCTCGAAATAGCTTCATTTTGCCTAGTGAATTTAGAA 660
 DB |||||

QY 601 ACCTTCGCTTTTGTGTCAGAGCTCGAAATAGCTTCATTTTGCCTAGTGAATTTAGAA 660
 DB |||||

QY 661 ATTATPCCCTTGAATACCTTGATCTTTTGGAAATATCTTTGAAACCAACAAAGTCTTTC 720
 DB |||||

QY 661 ATTATPCCCTTGAATACCTTGATCTTTTGGAAATATCTTTGAAACCAACAAAGTCTTTC 720
 DB |||||

QY 721 CAGTAATAAGCTGCAAGCAACCAATTAATTTATGAACTTCTGCAAGCAACCAATTAAT 780
 DB |||||

QY 721 CAGTAATAAGCTGCAAGCAACCAATTAATTTATGAACTTCTGCAAGCAACCAATTAAT 780
 DB |||||

QY 781 ATAATAGGATTCATATGCTCTCATATCAATTCATCTCCAGATTTTCGATA 840
 DB |||||

QY 781 ATAATAGGATTCATATGCTCTCATATCAATTCATCTCCAGATTTTCGATA 840
 DB |||||

QY 841 CCGCAAAAATTTGTGTGTGGAGATTTCTGTCTGAACCTTTTCAATTCAGGAAGCACTACTA 900
 DB |||||

QY 841 CCGCAAAAATTTGTGTGTGGAGATTTCTGTCTGAACCTTTTCAATTCAGGAAGCACTACTA 900
 DB |||||

QY 901 CCATGATCTGCATCTCTGTGCCCACACTGTGCTTAGTAGAATAATTTGGGTGGTACTG 960
 DB |||||

DB 901 CCATGATCTGCATCTCTGTGCCCACACTGTGGTCTTAGTAGATAAATTTGGGTGGTACTG 960
 QY 961 AAGCACCTATTATCTCTTATTTCTGTTCTCTAGGCTGTATGTTAAATTCCTCTGATATGT 1020
 DB |||||

DB 961 AAGCACCTATTATCTCTTATTTCTGTTCTCTAGGCTGTATGTTAAATTCCTCTGATATGT 1020
 QY 1021 TAAAGTAATGGTGGAGACCAAGAAAGAAATTTCAATAACAGATCAGTTTGGGGTGGATG 1080
 DB |||||

DB 1021 TAAAGTAATGGTGGAGACCAAGAAAGAAATTTCAATAACAGATCAGTTTGGGGTGGATG 1080
 QY 1081 TATCATTTTTCAGGCTCAAAATTTGAGTAAGGAAGATTTCTGTATCTTCTGCTGGAGAGA 1140
 DB |||||

DB 1081 TATCATTTTTCAGGCTCAAAATTTGAGTAAGGAAGATTTCTGTATCTTCTGCTGGAGAGA 1140
 QY 1141 GGAATGTGTATAGTACTCATTTTAGATGACTCCAAAACCTTTTATAAACAATTTTAGT 1200
 DB |||||

DB 1141 GGAATGTGTATAGTACTCATTTTAGATGACTCCAAAACCTTTTATAAACAATTTTAGT 1200
 QY 1201 TTTAAAAAATAAAAAA 1215
 DB |||||

DB 1201 TTTAAAAAATAAAAAA 1215
 RESULT 2
 AAI59205
 ID AAI59205 standard; cDNA; 1471 BP.
 XX
 AC AAI59205;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 1408.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW ankyrotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-USC34263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYSB-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 XX WPI; 2001-442253/47.
 DR P-PSDB; AAM40049.
 XX
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 PS Claim 1; SEQ ID NO 1408; 10078pp; English.
 XX
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the

CC encoded polypeptides (AA38642-AA42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as; Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification

XX Sequence 1471 BP; 433 A; 305 C; 301 G; 432 T; 0 U; 0 Other;
 Query Match 94.9%; Score 1153.2; DB 4; Length 1471;
 Best Local Similarity 99.7%; Pred. No. 3.2e-290;
 Matches 1155; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 58 TTTTTCAGCCATTTCCAGCAGTTTAAAGAGTTTCTTTCAGCTATGACACTGGCTCATTA 117
 DB 314 TAAGTAAGGCCATTTCCAGCAGTTTAAAGAGTTTCTTTCAGCTATGACACTGGCTCATTA 373
 QY 118 GAGGCTGTAACTTGTATACACAGTTTCAACGCTCACACAGTGAAGACTTCAGAAATTG 177
 DB 374 GAGGCTGTAACTTGTATACACAGTTTCAACGCTCACACAGTGAAGACTTCAGAAATTG 433
 QY 178 AAAACTTTAAACTAAATGGTTATACATCCAAAGAGCTATCTCTTAAGTAAGAATT 237
 DB 434 AAAACTTTAAACTAAATGGTTATACATCCAAAGAGCTATCTCTTAAGTAAGAATT 493
 QY 238 TTCCATATCTCTTGGNACATCTTCAGACTTCTACTGTGGGCTTGTCCGAGTTGATATGC 297
 DB 494 TTCCATATCTCTTGGNACATCTTCAGACTTCTACTGTGGGCTTGTCCGAGTTGATATGC 553
 QY 298 GTATGCTTTGCTTAAAGAGCTTTAGGAATTAGACTTGAGTCACACACCATATAAAAAGC 357
 DB 554 GTATGCTTTGCTTAAAGAGCTTTAGGAATTAGACTTGAGTCACACACCATATAAAAAGC 613
 QY 358 TTCCAGCTACAAATGGAGACCTTCATACACTTCAAGAACTTAACCTGAAGTCAATCACT 417
 DB 614 TTCCAGCTACAAATGGAGACCTTCATACACTTCAAGAACTTAACCTGAAGTCAATCACT 673
 QY 418 TGGAGTCATTTAGTGTAGCTTGTGTCATCTTACACTCCAGAGTCACTTCGAGTTTGG 477
 DB 674 TGGAGTCATTTAGTGTAGCTTGTGTCATCTTACACTCCAGAGTCACTTCGAGTTTGG 733
 QY 478 ACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGCAAGTTTGGCCAGCTCCAGGAACCTTA 537
 DB 734 ACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGCAAGTTTGGCCAGCTCCAGGAACCTTA 793
 QY 538 AGAATTTAAACTTCAGCAATATGAATGATTCATATTTCTTGGAGATAGGACAACTTA 597
 DB 794 AGAATTTAAACTTCAGCAATATGAATGATTCATATTTCTTGGAGATAGGACAACTTA 853
 QY 598 TAAACCTTTGGCTTTTGTGTCAGCAGCTCGAATAAGCTTCCATTTTGGCTAGTGAATCTTA 657
 DB 854 TAAACCTTTGGCTTTTGTGTCAGCAGCTCGAATAAGCTTCCATTTTGGCTAGTGAATCTTA 913
 QY 658 GAAATTTATCCCTTGAATATCTTGGATCTTTTGGAAATACCTTTTGAACAAACCAAGTCC 717
 DB 914 GAAATTTATCCCTTGAATATCTTGGATCTTTTGGAAATACCTTTTGAACAAACCAAGTCC 973
 QY 718 TTCCAGTAAATAGCTTGCAGACCACTTAACCTTTATTTGGAATCTTCTGCACCAACCATAT 777
 DB 974 TTCCAGTAAATAGCTTGCAGACCACTTAACCTTTATTTGGAATCTTCTGCACCAACCATAT 1033
 QY 778 TACATAATAGGATTCATATGCTCTCATATCATTTCCATTCATCTCTGCCAAGATTGG 837
 DB 1034 TACATAATAGGATTCATATGCTCTCATATCATTTCCATTCATCTCTGCCAAGATTGG 1093

QY 838 ATACCGCAAAATTTGTGTTTGGGAAGATTCGTCTGAACCTCTTTCATTCAAGGAACCTA 897
 DB 1094 ATACCGCAAAATTTGTGTTTGGGAAGATTCGTCTGAACCTCTTTCATTCAAGGAACCTA 1153
 QY 898 CTACCATGAATCTGCATTTCTGTGTCGCCACACACGTGTCTTAGTAGATAAATTTGGGTGGTA 957
 DB 1154 CTACCATGAATCTGCATTTCTGTGTCGCCACACACGTGTCTTAGTAGATAAATTTGGGTGGTA 1213
 QY 958 CTGAGCACCTATATCTCTTATTTCTGTCTCTAGGCTGTATGTAAATTCCTCTGTGATA 1317
 DB 1214 CTGAGCACCTATATCTCTTATTTCTGTCTCTAGGCTGTATGTAAATTCCTCTGTGATA 1273
 QY 1018 TGTTTAAAGTAAATGGGTGAGACCCAGAAAAAGAAATTTCAATAACAGATCAGTTTGGGTGTC 1077
 DB 1274 TGTTTAAAGTAAATGGGTGAGACCCAGAAAAAGAAATTTCAATAACAGATCAGTTTGGGTGTC 1333
 QY 1078 ATGTATGATTTTGCAGCGTCAAAATTCGAGTAAGGGAAGATTTCTGTATATCTTCTGTGAGA 1137
 DB 1334 ATGTATGATTTTGCAGCGTCAAAATTCGAGTAAGGGAAGATTTCTGTATATCTTCTGTGAGA 1393
 QY 1138 GGAGGAATGTATATGTTACTCATTTTAGATGACTCCAAAACCTTTTATTAACCAATTTT 1197
 DB 1394 GGAGGAATGTATATGTTACTCATTTTAGATGACTCCAAAACCTTTTATTAACCAATTTT 1453
 QY 1198 AGTTTAAAAAATAAAAAA 1215
 DB 1454 AGTTTAAAAAATAAAAAA 1471

RESULT 3

ADB49189

ID ADB49189 standard; cDNA; 1471 BP.

XX ADB49189;

XX 04-DEC-2003 (first entry)

XX Novel human cDNA SEQ ID NO 1099.

XX ss; cancer; neurodegenerative disease; human.

XX Homo sapiens.

XX US2003104529-A1.

XX 05-JUN-2003.

XX 04-JAN-2002; 2002US-00037270.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 19-JUL-2000; 2000US-00620312.

XX (ZHOU/) ZHOU P.

XX (TANG/) TANG Y T.

XX (LIUC/) LIU C.

XX (ASUN/) ASUNDI V.

XX (DRMA/) DRMANAC R T.

XX Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;

XX WPI; 2003-678194/64.

XX New polynucleotide, useful for treating diseases e.g., cancer or neurodegenerative diseases.

XX Claim 1; SEQ ID NO 1099; 99pp; English.

XX The invention relates to a polynucleotide comprising a sequence given in the specification, or its mature protein-coding portion, or its complement. The polynucleotide is useful for treating diseases e.g., cancer or neurodegenerative diseases and many others listed in the specification. The present sequence represents a novel human cDNA. Note:

CC The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format directly from USPTO
 CC at seqdata.uspto.gov/sequence.html?DocID=20030104529.
 XX
 SQ Sequence 1471 BP; 433 A; 305 C; 301 G; 432 T; 0 U; 0 Other;
 Query Match 94.9%; Score 1153.2; DB 8; Length 1471;
 Best Local Similarity 99.7%; Pred. No. 3.7e-290;
 Matches 1155; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 58 TTATTCAGGCAATTCAGCAGATTTAAAGGTTTCCCTTTGAGCTATGAGACTGGCTCATTA 117
 DB 314 TAAGTAAGGCAATTCAGCAGATTTAAAGGTTTCCCTTTGAGCTATGAGACTGGCTCATTA 373
 QY 118 GAGGCTGTAATGTTGATACACCACTTTCAACGCTCACACCACTGAAGACTTTCAGAAATTTG 177
 DB 374 GAGGCTGTAATGTTGATACACCACTTTCAACGCTCACACCACTGAAGACTTTCAGAAATTTG 433
 QY 178 AAAACCTTTAAACCTAAATAGTATATCAATCAACCAAAAGACTATCCTCTAAGTAAGAATT 237
 DB 434 AAAACCTTTAAACCTAAATAGTATATCAATCAACCAAAAGACTATCCTCTAAGTAAGAATT 493
 QY 238 TTCCATATTCCTTGGACATCTTCAGACTTCTTACTGTGGCTTGTGGAGTTGATATGC 297
 DB 494 TTCCATATTCCTTGGACATCTTCAGACTTCTTACTGTGGCTTGTGGAGTTGATATGC 553
 QY 298 GTATGCTTTGCTTTAAAGGCTTTAGGAATTTAGACTTTGAGTCACACCAATATAAAAGC 357
 DB 554 GTATGCTTTGCTTTAAAGGCTTTAGGAATTTAGACTTTGAGTCACACCAATATAAAAGC 613
 QY 358 TTCAGCTACAAATGGAGACCTCATACACCTTCAAGAACTTAACCTGAATGACAACTCACT 417
 DB 614 TTCAGCTACAAATGGAGACCTCATACACCTTCAAGAACTTAACCTGAATGACAACTCACT 673
 QY 418 TGGAGTCAATTTAGTGTAGCTTGTGTCTATCTACCTCCAGAACTCACTTCGGAGTTGG 477
 DB 674 TGGAGTCAATTTAGTGTAGCTTGTGTCTATCTACCTCCAGAACTCACTTCGGAGTTGG 733
 QY 478 ACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGCAAGTTTGTGCACTCCAGCAACTTA 537
 DB 734 ACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGCAAGTTTGTGCACTCCAGCAACTTA 793
 QY 538 AGAATTTAAACCTTGAGTAATGAATGATTAATTTCTTGGCAAGATAGGACAACTAA 597
 DB 794 AGAATTTAAACCTTGAGTAATGAATGATTAATTTCTTGGCAAGATAGGACAACTAA 853
 QY 598 TAAACCTTCGCTTTTGTGACGAGCTGAAATGAGCTTCCATTTTCCCTAGTGAATTTA 657
 DB 854 TAAACCTTCGCTTTTGTGACGAGCTGAAATGAGCTTCCATTTTCCCTAGTGAATTTA 913
 QY 658 GAAATTTATCCCTTGAATATCTGGATCTTTTGGAAATATCTTTTGAACCAACCAAAAGTCC 717
 DB 914 GAAATTTATCCCTTGAATATCTGGATCTTTTGGAAATATCTTTTGAACCAACCAAAAGTCC 973
 QY 718 TTCCAGTAATAGCTTCAGACCACTTAATTTATTTGGATCTTCTGCAAGAACCAATAT 777
 DB 974 TTCCAGTAATAGCTTCAGACCACTTAATTTATTTGGATCTTCTGCAAGAACCAATAT 1033
 QY 778 TACATAATAGGATTCATATGGCTCTCATATCATCTCCATTTCCATCTCTGCAAGATTTGG 837
 DB 1034 TACATAATAGGATTCATATGGCTCTCATATCATCTCCATTTCCATCTCTGCAAGATTTGG 1093
 QY 838 ATACCGCAAAATTTGTTTGGAGATTCCTGTCTGAATCTTTCAATTCAGGAACTA 897
 DB 1094 ATACCGCAAAATTTGTTTGGAGATTCCTGTCTGAATCTTTCAATTCAGGAACTA 1153
 QY 898 CTACCATGATCTGCATCTGTTGCCACACTGTGCTCTTAGTAGATAATTTGGTGGA 957
 DB 1154 CTACCATGATCTGCATCTGTTGCCACACTGTGCTCTTAGTAGATAATTTGGTGGA 1213
 QY 958 CTGAAGCACTTATCTCTTATTTCTTCTCTAGGCTGTATGTAATTTCTCTGTATA 1017
 DB 1214 CTGAAGCACTTATCTCTTATTTCTTCTCTAGGCTGTATGTAATTTCTCTGTATA 1273

QY 1018 TGTTAAAGTAATGGTGGACACGAGAAAGAAATTTCAATAACAGATCAGTTTGGGGTGC 1077
 DB 1274 TGTTAAAGTAATGGTGGACACGAGAAAGAAATTTCAATAACAGATCAGTTTGGGGTGC 1333
 QY 1078 ATGTATGATTTTGCAGCGCTCAAAATGGAGTAAGGAAGATTTCTGTATATCTGCTGGAGA 1137
 DB 1334 ATGTATGATTTTGCAGCGCTCAAAATGGAGTAAGGAAGATTTCTGTATATCTGCTGGAGA 1393
 QY 1138 GGAGGAATGTATAGTTACTCTCAATTTAGATGATCTCCAAACCTTTTATTTAAACCAATTTT 1197
 DB 1394 GGAGGAATGTATAGTTACTCTCAATTTAGATGATCTCCAAACCTTTTATTTAAACCAATTTT 1453
 QY 1198 AGTTTAAAAAATAAAA 1215
 DB 1454 AGTTTAAAAAATAAAA 1471

RESULT 4
 ABK90911
 ID ABK90911 standard; cdna; 1600 BP.
 XX
 AC ABK90911;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE cdna encoding fruit fly LRR47 polypeptide 47-33.88.
 XX
 KW Fruit fly; LRR47 polypeptide 47-33.88; embryonic development deformity;
 KW tumour; diabetes; menstrual disorder; peptide ulcer; arrhythmia; anaemia;
 KW epilepsy; gene; ss.
 XX
 OS Drosophila sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 462..1388
 FT /*tag= a
 FT /product= "Fruit fly LRR47 polypeptide 47-33.88"
 XX
 CN1341640-A.
 XX
 PD 27-MAR-2002.
 XX
 PF 05-SEP-2000; 2000CN-00125025.
 XX
 PR 05-SEP-2000; 2000CN-00125025.
 XX
 PA (SHAN-) SHANGHAI BIODOOR GENE DEV CO LTD.
 XX
 PI Mao Y, Xie Y;
 XX
 DR MPI; 2002-520716/56.
 DR P-PSDB; ABG31598.
 XX
 PT A fruit fly LRR47 polypeptide 47-33.88, useful for curing e.g. tumors and
 XX diabetes.
 PS Claim 6; Page 25-26 (Disclosure); 33pp; Chinese.
 XX
 CC The present invention relates to a new fruit fly LRR47 polypeptide 47-
 CC 33.88. The polypeptide is useful for curing several diseases, such as
 CC embryonic development deformity, tumour, diabetes, menstrual disorder,
 CC peptide ulcer, arrhythmia, anaemia and epilepsy. The present nucleic acid
 CC sequence encodes the fruit fly LRR47 polypeptide 47-33.88 of the
 CC invention
 XX
 SQ Sequence 1600 BP; 455 A; 336 C; 349 G; 460 T; 0 U; 0 Other;
 Query Match 94.9%; Score 1153; DB 6; Length 1603;
 Best Local Similarity 100.0%; Pred. No. 3.7e-290;
 Matches 1153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 CAGGCCATTTCCAGCAGTGTAAAGGTTTCCTTTCAGCTATGAGACTGGCTCATAGAGC 122

Db 423 CAGGCCATTTCCAGCGATTTAAAGGTTTCCITTCAGCTATGAGACTGGCTCATAGGC 482
QY |||||||
Db 123 TCTAATGTTGATACACCGATTCAAGCTTCACACGCTGAGAGCTTCAGAAATTGAAAC 182
QY |||||||
Db 483 TGTAAATGTTGATACACCGATTCAAGCTTCACACGCTGAGAGCTTCAGAAATTGAAAC 542
QY |||||||
Db 183 TTTAAACCTAAATGTTATCATCTCAACCAAAAGACTATCCTCTAAGTAAAGATTTTCCA 242
QY |||||||
Db 543 TTTAAACCTAAATGTTATCATCTCAACCAAAAGACTATCCTCTAAGTAAAGATTTTCCA 602
QY |||||||
Db 243 TATTCTCTGGAACATCTTCAGACTCTTCTACTGTGGGCTTTGTCGAGTTGATATGCGTATG 302
QY |||||||
Db 603 TATTCTCTGGAACATCTTCAGACTCTTCTACTGTGGGCTTTGTCGAGTTGATATGCGTATG 662
QY |||||||
Db 303 CTTTGTCTTAAAGCCTTAGGAAATTAGACTTGAGTCAACACCATATATAAAAGCTTCCA 362
QY |||||||
Db 663 CTTTGTCTTAAAGCCTTAGGAAATTAGACTTGAGTCAACACCATATATAAAAGCTTCCA 722
QY |||||||
Db 363 GCTACAATTTGGAGACCTCATACACCTTCAGACTTAACTGTAATCAACATCACTTGGAG 422
QY |||||||
Db 723 GCTACAATTTGGAGACCTCATACACCTTCAGACTTAACTGTAATCAACATCACTTGGAG 782
QY |||||||
Db 423 TCATTAGTGTAGCTTGTGTCTCAATTCACACTCCAGAGTCACTTCGGAGTTTGGACCTC 482
QY |||||||
Db 783 TCATTAGTGTAGCTTGTGTCTCAATTCACACTCCAGAGTCACTTCGGAGTTTGGACCTC 842
QY |||||||
Db 483 AGCAGAACAAATCAAGGACTCCCTGTGGAGTTTGGAGCTCCAGGAACTTAAGAAAT 542
QY |||||||
Db 843 AGCAGAACAAATCAAGGACTCCCTGTGGAGTTTGGAGCTCCAGGAACTTAAGAAAT 902
QY |||||||
Db 543 TTAATAACTGAGATTAATGAATGATTCATTTCCCTGCAAGATAGACAACTTAATAAC 602
QY |||||||
Db 903 TTAATAACTGAGATTAATGAATGATTCATTTCCCTGCAAGATAGACAACTTAATAAC 962
QY |||||||
Db 603 CTTCTCTTTTGTGAGCAGCTCGAAATAGCTTTCCATTTTTCCTAGTGAATTTAGAAAT 662
QY |||||||
Db 963 CTTCTCTTTTGTGAGCAGCTCGAAATAGCTTTCCATTTTTCCTAGTGAATTTAGAAAT 1022
QY |||||||
Db 663 TTATCCCTTGATATCTTGATCTTTTGGAAATACCTTTTGAACAACTTTCCTTCCA 722
QY |||||||
Db 1023 TTATCCCTTGATATCTTGATCTTTTGGAAATACCTTTTGAACAACTTTCCTTCCA 1082
QY |||||||
Db 723 GTAATAAGCTCAAGCAGCACTTAATCACTTATTGGAAATCTTTCGACAGCACTTATACAT 782
QY |||||||
Db 1083 GTAATAAGCTCAAGCAGCACTTAATCACTTATTGGAAATCTTTCGACAGCACTTATACAT 1142
QY |||||||
Db 783 AATAGGATTCATATGGCTCTCATATCATCTCCATTCATCTCGCCAGATTTGGATACC 842
QY |||||||
Db 1143 AATAGGATTCATATGGCTCTCATATCATCTCCATTCATCTCGCCAGATTTGGATACC 1202
QY |||||||
Db 843 GCAAAAATTTGGTTTGGAGATCTCTGTAACCTTTTCAATTCAGGAACTATCTACC 902
QY |||||||
Db 1203 GCAAAAATTTGGTTTGGAGATCTCTGTAACCTTTTCAATTCAGGAACTATCTACC 1262
QY |||||||
Db 903 ATGAATCTGCATCTCTGTGCCACACTGTGGTCTTAGTAGATAAATTTGGGTGCTAGAA 962
QY |||||||
Db 1263 ATGAATCTGCATCTCTGTGCCACACTGTGGTCTTAGTAGATAAATTTGGGTGCTAGAA 1322
QY |||||||
Db 963 GCACCTATTTCTATTTCTGTCTTAGCTGTATGTTAAATTCCTCTGATATGTTA 1022
QY |||||||
Db 1323 GCACCTATTTCTATTTCTGTCTTAGCTGTATGTTAAATTCCTCTGATATGTTA 1382
QY |||||||
Db 1023 AAGTAATGGGTGAGACCAAGAAAGAAATTTCAATAACAGATCAGTTTGGGGTGACGTA 1082
QY |||||||
Db 1383 AAGTAATGGGTGAGACCAAGAAAGAAATTTCAATAACAGATCAGTTTGGGGTGACGTA 1442
QY |||||||
Db 1083 TGATTTTGCAGCGTCAAAATGAGTAAAGGAGATTTCTGTATCTTGTCTGGAGAGGAG 1142
QY |||||||
Db 1443 TGATTTTGCAGCGTCAAAATGAGTAAAGGAGATTTCTGTATCTTGTCTGGAGAGGAG 1502
QY |||||||
Db 1143 RATGTGATAGTACTCATTTAGATGACTCCAAACTTTATTAACCAATTTAGTTT 1202
QY |||||||

Db 1503 AATGTGTATAGTTACTCATTTAGATGACTCCAAAACCTTTATTAAAAACCAATTTTACTTT 1562
QY 1203 TAAAAAATAAAAA 1215
Db 1563 TAAAAAATAAAAA 1575
RESULT 5
AAS33220
ID AAS33220 standard; cDNA; 1341 BP.
XX AAS33220;
AC AC
XX DT
XX 04-DEC-2001 (first entry)
DE DNA encoding human secreted protein, Seq ID No 179.
XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
KW rheumatoid arthritis; antiarteriosclerotic; cardiac; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
KW Gaucher's disease; neurological disease; cerebrovascular disorder;
KW thrombosis; wound healing; ss.
XX Homo sapiens.
XX OS
XX WO200153326-A2.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001347.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225274P.
PR 14-AUG-2000; 2000US-0225275P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.

Qy 118 GAGGCTGTAATGTTGATACACCAAGTTTCAACGCTCACACCCAGTGAGACTTCAGAAATTG 177

[illegible]

Db	1320	AGTTTAAAAA	1336
RESULT 6			
AAI60991			
ID	AAI60991	standard; cDNA; 1287 BP.	
XX	AAI60991;		
XX	22-OCT-2001	(first entry)	
XX	Human polynucleotide SEQ ID NO 4983.		
XX	Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;		
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;		
KW	leukaemia; ss.		
XX			
OS	Homo sapiens.		
XX			
FN	WO200153312-A1.		
XX			
PD	26-JUL-2001.		
XX			
PF	26-DEC-2000; 2000WO-US034263.		
XX			
PR	23-DEC-1999; 99US-00471275.		
PR	21-JAN-2000; 2000US-00488725.		
PR	25-APR-2000; 2000US-0052317.		
PR	20-JUN-2000; 2000US-00598042.		
PR	19-JUL-2000; 2000US-00620312.		
PR	03-AUG-2000; 2000US-00653450.		
PR	14-SEP-2000; 2000US-00662191.		
PR	19-OCT-2000; 2000US-00693036.		
PR	29-NOV-2000; 2000US-00727344.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Tang YT, Liu C, Asundi V, Chen R, Xia Y, Qian XB, Ren F, Wang D;		
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;		
PI	Zhou F, Goodrich R, Drmanac RT;		
XX			
DR	WPI; 2001-442253/47.		
DR	P-PSDB; AAM41835.		
XX			
PT	Novel nucleic acids and polypeptides, useful for treating disorders such		
PT	as central nervous system injuries.		
XX			
PS	Claim 1; SEQ ID NO 4980; 10079pp; English.		
XX			
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and the		
CC	encoded polypeptides (AAM38642-AAM42213) with nontropic.		
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful		
CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	of the invention may be used to treat diseases of the peripheral nervous		
CC	system, such as peripheral nervous injuries, peripheral neuropathy and		
CC	centralised neuropathies and central nervous system diseases, such as		
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic		
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the		
CC	utilisation of the activities such as: Immune system suppression.		
CC	Activation/inhibit activity, chemoractic/chemokinetic activity, haemostatic		
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,		
CC	assays for receptor activity, arthritis and inflammation, leukaemias and		
CC	C.N.S disorders. Note: The sequence data for this patent did not form		
CC	part of the printed specification		
XX			
SQ	Sequence 1287 BP; 363 A; 282 C; 264 G; 378 T; 0 U; 0 Other;		
Query Match	78.3%;	Score 951.4;	DB 4; Length 1287;
Best Local Similarity	99.4%;	Pred. No. 1.2e-237;	
Matches 955; Conservative	0;	Mismatches 6;	Indels 0; Gaps 0;

QY 58 TTATTGAGGCAATTCAGCAGCTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCATTA 117
DB 326 TAAGTAAGGCAATTCAGCAGCTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCATTA 385
QY 118 GAGGCTGTAATGTTGATACACCAAGTTTCAACGCTCACACAGTGAAGACTTCAGATTTCG 177
DB 386 GAGGCTGTAATGTTGATACACCAAGTTTCAACGCTCACACAGTGAAGACTTCAGATTTCG 445
QY 178 AAACTTTTAAATCAAAATGTTTATCATCATCCAAAAGACTATCCCTTAAGTAAGAAATT 237
DB 446 AAACTTTTAAATCAAAATGTTTATCATCATCCAAAAGACTATCCCTTAAGTAAGAAATT 505
QY 238 TTCCATATTCCTTGGACATCTTCAGACTTCCTACTGTGGCTTGTCCGAGTTGATATGC 297
DB 506 TTCCATATTCCTTGGACATCTTCAGACTTCCTACTGTGGCTTGTCCGAGTTGATATGC 565
QY 298 GTATGCTTTGCTTAAAGGCTTTAGGAAATTTAGACTTGGTCAACCATATATAAAGC 357
DB 566 GTATGCTTTGCTTAAAGGCTTTAGGAAATTTAGACTTGGTCAACCATATATAAAGC 625
QY 358 TTCCAGCTCAATTTGGAGACTCATACCTTCAAGAACTTAACTGAATGACAACTACT 417
DB 626 TTCCAGCTCAATTTGGAGACTCATACCTTCAAGAACTTAACTGAATGACAACTACT 685
QY 418 TGGAGTCAATTTAGTGTAGCTTGTGTCATCTTCTACACTCCAGAACTCACTCGGAGTTTGG 477
DB 686 TGGAGTCAATTTAGTGTAGCTTGTGTCATCTTCTACACTCCAGAACTCACTCGGAGTTTGG 745
QY 478 ACCTCAGCAAGCAAAATCAAGGCACTCCCTGTGCAAGTTTTCGCACTCCAGCAACTTA 537
DB 746 ACCTCAGCAAGCAAAATCAAGGCACTCCCTGTGCAAGTTTTCGCACTCCAGCAACTTA 805
QY 538 AGAATTTAAACTGTGACGTAATGAATTTGAACTTCAATTCCTTGCAGATAGGCAACTAA 597
DB 806 AGAATTTAAACTGTGACGTAATGAATTTGAACTTCAATTCCTTGCAGATAGGCAACTAA 865
QY 598 TAAACCTTCGCTTTTGTGACGAGCTCGAATAAGCTTCCATTTTGCCTGAGTGAATTA 657
DB 866 TAAACCTTCGCTTTTGTGACGAGCTCGAATAAGCTTCCATTTTGCCTGAGTGAATTA 925
QY 658 GAAATTTATCCCTTGAATCTTGGATCTTTTGGAAATACCTTTTGAACCAACCAAGTCC 717
DB 926 GAAATTTATCCCTTGAATCTTGGATCTTTTGGAAATACCTTTTGAACCAACCAAGTCC 985
QY 718 TTCCAGTATTAAGCTGCAAGCACTTAACTTTATTTGGAATCTTTCGACGCAACCATAT 777
DB 986 TTCCAGTATTAAGCTGCAAGCACTTAACTTTATTTGGAATCTTTCGACGCAACCATAT 1045
QY 778 TACATATAGGATTCATATGCTCTCATATCATTCATTCATTCATTCATTCATTCATTCAT 837
DB 1046 TACATATAGGATTCATATGCTCTCATATCATTCATTCATTCATTCATTCATTCATTCAT 1105
QY 838 ATACCCCAAAATTTGTTTGGAGATCTTGTCTGAATCTTTCATTCAGGAACCTA 897
DB 1106 ATACCCCAAAATTTGTTTGGAGATCTTGTCTGAATCTTTCATTCAGGAACCTA 1165
QY 898 CTACCATGATCTGATCTGTTGCCACACTGTTGCTTAGTAGATAAATTCGGTGGTA 957
DB 1166 CTACCATGATCTGATCTGTTGCCACACTGTTGCTTAGTAGATAAATTCGGTGGTA 1225
QY 958 CTGAAGCACTTATCTCTTATTTCTGTTCTCTAGGCTGTTATGTTAACTCTCTGATA 1017
DB 1226 CTGAAGCACTTATCTCTTATTTCTGTTCTCTAGGCTGTTATGTTAACTCTCTGATA 1285
QY 1018 T 1018
DB 1286 T 1286

RESULT 7
AAD27818
ID AAD27818 standard; DNA; 1251 BP.

XX AAD27818;
AC 18-APR-2002 (first entry)
DT Human leucine rich repeat protein gene, sbg442358PROA #2.
XX
DE Human; therapy; wound healing disorder; vaccine; cancer; infection;
XX autoimmune disorder; haematopoietic disorder; inflammation; arthritis;
KW Parkinson's disease; Huntington's chorea; schizophrenia; antiarrhythmic;
KW multiple sclerosis; Alzheimer's disease; analgesic; cardiac; asthma;
KW ischaemia; stroke; AIDS; bone disease; atherosclerosis; brain disorder;
KW depression; cardiovascular disease; myocardial infarction; renal failure;
KW respiratory disease; liver disorder; Fanconi's syndrome; spleen disorder;
KW type II diabetes mellitus; skeletal muscle disorder; immunosuppressive;
KW hyperplasia; renal disease; hypoglycaemia; gastrointestinal disease;
KW neoplastic; cirrhosis; Hodgkin's disease; neuroleptic; antiinflammatory;
KW haemostatic; vulvovaginitis; anticonvulsant; antirheumatic; neuroprotective;
KW nephrotropic; hypotensive; vasotropic; cytostatic; cerebroprotective;
KW allergy; leucine rich repeat protein; gene; ds.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH 1. .1251
FT /*tag= a
FT /product= "Human leucine rich repeat protein"
XX
XX WO200198342-A1.
XX
XX 27-DEC-2001.
XX
XX 22-JUN-2001; 2001WO-US019929.
XX
XX 22-JUN-2000; 2000US-0213156P.
XX
XX 22-JUN-2000; 2000US-0213161P.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX (SMIK) SMITHKLINE BEECHAM PLC.
XX (GLAX) GLAXO GROUP LTD.
XX
XX Agarwal P, Cogswell JP, Kabnic KS, Lai Y, Martensen SA,
PI Murdock PR, Smith RF, Strum JC, Xiang Z, Xie Q, Rizni SK;
XX
XX WPI: 2002-139783/18.
XX P-PSDB; AAE17323.
XX
XX Novel secreted and membrane-associated polypeptides and polynucleotides
XX useful for preventing, ameliorating or correcting dysfunction or disease
XX including diabetes, cancer, hypertension and growth abnormalities.
XX
XX Claim 2; Page 99; 138pp; English.
XX
XX The invention relates to secreted and membrane-associated polypeptides
XX and polynucleotides. The sequences of the invention are useful in
XX diagnostic assays for detecting diseases associated with inappropriate
XX activity or levels of these polynucleotides, and in identifying their
XX agonists and antagonists that are potentially useful in therapy. The
XX sequences of the invention are useful as vaccines for inducing
XX immunological response. The sequences of the invention are useful for
XX treating cancers, infections, autoimmune disorders, haematopoietic
XX disorders, wound healing disorders, cholesterol ester storage disease,
XX inflammation, congenital muscular dystrophy, junctional epidermolysis
XX bullosa, Parkinson's disease, Huntington's chorea, multiple sclerosis,
XX viral and bacterial infections, Alzheimer's disease, asthma, arthritis,
XX allergies, schizophrenia, sbg44245PROA-associated disorders,
XX septicemia, psoriasis, inflammatory bowel disease, transplant rejection,
XX graft versus host disease, ischaemia, stroke, acute respiratory disease
XX syndrome, restenosis, brain injury, AIDS, bone diseases, atherosclerosis,
XX brain disorders including parasupranuclear palsy, myotonic dystrophy,
XX depression, anxiety disorders and sleep disorders, cardiovascular
XX diseases including congestive heart failure and myocardial infarction,
XX respiratory diseases including chronic obstructive pulmonary disease,

CC acute bronchitis and adult respiratory distress syndrome, liver disorders
CC including hypercholesterolaemia, hypertriglyceridaemia, cirrhosis, viral
CC and non-viral hepatitis, type II diabetes mellitus, renal disease
CC syndrome, acute and chronic renal failure, glomerulonephritis, Fanconi's
CC disease, cystinuria, skeletal muscle disorders including hypoglycaemia
CC and tendinitis, gastrointestinal diseases including intestinal
CC obstruction and tropical sprue, spleen disorders including hypersplenism,
CC Hodgkin's disease and malignant lymphoma, testicular cancer, male
CC reproductive diseases including low testosterone and male infertility.
CC The present sequence is human leucine rich repeat protein gene
XX
SQ Sequence 1251 BP; 358 A; 275 C; 248 G; 370 T; 0 U; 0 Other;

Query Match 78.2%; Score 950.2; DB 6; Length 1251;
Best Local Similarity 99.1%; Pred. No. 2.4e-237;
Matches 968; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

QY 58 TTATTCAGGCCAATTCAGCAGTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCATTA 117
DB 275 TAAGTAAGGCCAATTCAGCAGTTTAAAGGTTTCCTTCAGCTATGAGACTGGCTCATTA 334
QY 118 GAGGCTGTATGTGATACACAGTTTCAACGCTCACACAGTGRAGACTTCAGAAATTG 177
DB 335 GAGGCTGTATGTGATACACAGTTTCAACGCTCACACAGTGRAGACTTCAGAAATTG 394
QY 178 AAAACCTTTAAACCTAAATGGTTTATCACATCCAAAAAGACTATCCTCTAAAGTAAGAATT 237
DB 395 AAAACCTTTAAACCTAAATGGTTTATCACATCCAAAAAGACTATCCTCTAAAGTAAGAATT 454
QY 238 TTCCATTTCTTGGAACTCTTCAGACTTCTTACTGTGGGCTTGTCCGAGTTGATATGC 297
DB 455 TTCCATTTCTTGGAACTCTTCAGACTTCTTACTGTGGGCTTGTCCGAGTTGATATGC 514
QY 298 GTATGCTTTGCTTAAAGGCTTAGAATTTAGACTGTAGTGTACACCACTATATAAAGG 357
DB 515 GTATGCTTTGCTTAAAGGCTTAGAATTTAGACTGTAGTGTACACCACTATATAAAGG 574
QY 358 TTCCAGCTACAAATTGAGACCTCATACACCTTCAAGAACTTAACCTGAATGACAACTCACT 417
DB 575 TTCCAGCTACAAATTGAGACCTCATACACCTTCAAGAACTTAACCTGAATGACAACTCACT 634
QY 418 TGGAGTCATTTAGTGTAGCTTGTGTCATTTACACTCCAGAAAGTCACTTCGGAGTTGG 477
DB 635 TGGAGTCATTTAGTGTAGCTTGTGTCATTTACACTCCAGAAAGTCACTTCGGAGTTGG 694
QY 478 ACCTCAGCAAGCAAAAAATCAAGGCACTCCCTGTGCAGTTTTCGACGCTCCAGAACTTA 537
DB 695 ACCTCAGCAAGCAAAAAATCAAGGCACTCCCTGTGCAGTTTTCGACGCTCCAGAACTTA 754
QY 538 AGAATTTAAACCTTGCAGATAATGAATGATTCATTTCTTCGCAAGATAGGACAACTAA 597
DB 755 AGAATTTAAACCTTGCAGATAATGAATGATTCATTTCTTCGCAAGATAGGACAACTAA 814
QY 598 TAAACCTTCGCTTTTGTGAGGCTGCAATTAAGCTTCCATTTTTCCTAGTGAATTTA 657
DB 815 TAAACCTTCGCTTTTGTGAGGCTGCAATTAAGCTTCCATTTTTCCTAGTGAATTTA 874
QY 658 GAAATTTATCCCTTGAATACCTTGGATCTTTTGGAAATATCTTTTGAAACCAACAAAGTCC 717
DB 875 GAAATTTATCCCTTGAATACCTTGGATCTTTTGGAAATATCTTTTGAAACCAACAAAGTCC 934
QY 718 TTCCAGTAATAAGCTGCAAGCAACCAATTAATCTTTATTTGAATCTTCTGCACGACCATAT 777
DB 935 TTCCAGTAATAAGCTGCAAGCAACCAATTAATCTTTATTTGAATCTTCTGCACGACCATAT 994
QY 778 TACAT-----AATAGGATTCATATGCTTCATATCAATTCATCTTCGCGAAG 831
DB 995 TACATTAATAGGAATAGGATTCATATGCTTCATATCAATTCATCTTCGCGAAG 1054
QY 832 ATTTGGATACCGCAAAAAATTTGTGTTGTGGAGATTCCTGTCTGAATCTTTTCATTTCAAG 891
DB 1055 ATTTGGATACCGCAAAAAATTTGTGTTGTGGAGATTCCTGTCTGAATCTTTTCATTTCAAG 1114

QY 892 GAACTACTACCATGAATCTGATCTCTGTTGCCACACACTGTGGTCTTAGTAGATAATTTGG 951
DB 1115 GAACTACTACCATGAATCTGATCTCTGTTGCCACACACTGTGGTCTTAGTAGATAATTTGG 1174
QY 952 GTCGTACTAGGACCACTTATCTTATTTCTGTTCTTCTAGGCTGTATGTTAATTCCT 1011
DB 1175 GTCGTACTAGGACCACTTATCTTATTTCTGTTCTTCTAGGCTGTATGTTAATTCCT 1234
QY 1012 CTCGATATGTTAAAGTAA 1028
DB 1235 CTCGATATGTTAAAGTAA 1251

RESULT 8
AAD27817

ID AAD27817 standard; DNA; 780 BP.

XX

AC

AD27817;

DT 18-APR-2002 (first entry)

XX

DE Human leucine rich repeat protein gene, sbg442358PROa #1.

XX

Human; therapy; wound healing disorder; vaccine; cancer; infection;
autoimmune disorder; haematopoietic disorder; inflammation; arthritis;
Parkinson's disease; Huntington's chorea; schizophrenia; antiarrhythmic;
multiple sclerosis; Alzheimer's disease; analgesic; cardiac; asthma;
ischemia; stroke; AIDS; bone disease; atherosclerosis; brain disorder;
depression; cardiovascular disease; myocardial infarction; renal failure;
respiratory disease; liver disorder; Fanconi's syndrome; spleen disorder;
type II diabetes mellitus; skeletal muscle disorder; immunosuppressive;
hyperplasia; renal disease; hypoglycaemia; gastrointestinal disease;
neurotic; cirrhosis; Hodgkin's disease; neuroleptic; antiinflammatory;
haemostatic; vulvular; anticonvulsant; antineoplastic; neuroprotective;
nephrotropic; hypotensive; vasotropic; cytostatic; cerebroprotective;
allergy; leucine rich repeat protein; gene; ds.

XX

OS Homo sapiens.

XX

Key Location/Qualifiers

CDS

1..780

/*tag= a

/product= "Human leucine rich repeat protein"

XX

WO200198342-A1.

XX

27-DEC-2001.

XX

22-JUN-2001; 2001WO-US019929.

XX

22-JUN-2000; 2000US-0213156P.

XX

22-JUN-2000; 2000US-0213151P.

XX

(SMIK) SMITHKLINE BEECHAM CORP.

PA (SMIK) SMITHKLINE BEECHAM PLC.

PA (GLAX) GLAXO GROUP LTD.

XX

Agarwal P, Cogswell JP, Kabnick KS, Lai Y, Martensen SA,
Yardock PR, Smith RE, Strum JC, Xiang Z, Xie Q, Rizni SK,
P-PSDB; AAE17322.

DR WPI; 2002-139783/18.

XX

Novel secreted and membrane-associated polypeptides and polynucleotides
useful for preventing, ameliorating or correcting dysfunction or disease
including diabetes, cancer, hypertension and growth abnormalities.

PT

PT

XX

Claim 2; Page 98; 138pp; English.

XX

The invention relates to secreted and membrane-associated polypeptides
and polynucleotides. The sequences of the invention are useful in
diagnostic assays for detecting diseases associated with inappropriate
activity or levels of these polynucleotides, and in identifying their

agonists and antagonists that are potentially useful in therapy. The sequences of the invention are useful as vaccines for inducing an immunological response. The sequences of the invention are useful for treating cancers, infections, autoimmune disorders, hematopoietic disorders, wound healing disorders, cholesteryl ester storage disease, inflammation, congenital muscular dystrophy, junctional epidermolysis bullosa, Parkinson's disease, Huntington's chorea, multiple sclerosis, viral and bacterial infections, Alzheimer's disease, asthma, arthritis, allergies, schizophrenia, sbg44245PKOa-associated disorders, septicemia, psoriasis, inflammatory bowel disease, transplant rejection, graft versus host disease, ischemia, stroke, acute respiratory disease, syndrome, restenosis, brain injury, AIDS, bone diseases, atherosclerosis, brain disorders including paraneoplastic palsy, myotonic dystrophy, depression, anxiety disorders and sleep disorders, cardiovascular diseases including congestive heart failure and myocardial infarction, respiratory diseases including chronic obstructive pulmonary disease, acute bronchitis and adult respiratory distress syndrome, liver disorders including hypercholesterolemia, hypertriglyceridemia, cirrhosis, viral and non-viral hepatitis, type II diabetes mellitus, renal disease including acute and chronic renal failure, glomerulonephritis, Fanconi's syndrome, cystinuria, skeletal muscle disorders including hypoglycemia and tendinitis, gastrointestinal diseases including intestinal obstruction and tropical sprue, spleen disorders including hypersplenism, Hodgkin's disease and malignant lymphoma, testicular cancer, male reproductive diseases including low testosterone and male infertility. The present sequence is human leucine rich repeat protein gene

Sequence 780 BP; 250 A; 160 C; 125 G; 245 T; 0 U; 0 Other;

Query Match 64.0%; Score 777; DB 6; Length 780;
 Best Local Similarity 100.0%; Pred. No. 2.8e-192;
 Matches 777; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

12 ATGTATGATATATCTCCAGTGGAAATTTATATCTACAACTTTTATTCAGGCCATT 71
 1 ATGTATGATATATCTCCAGTGGAAATTTATATCTACAACTTTTATTCAGGCCATT 60
 72 TCAGAGAGTTTAAAGGTTTCTTTTCAGTATGAGCTGGCTCATAGAGCTGAATGTT 131
 61 TCAGAGAGTTTAAAGGTTTCTTTTCAGTATGAGCTGGCTCATAGAGCTGAATGTT 120
 132 GATACACAGTTTCAACGCTCACACAGTGAAGACTTCAGAAATTTGAAAACCT 191
 121 GATACACAGTTTCAACGCTCACACAGTGAAGACTTCAGAAATTTGAAAACCT 180
 192 AAAATGTTATCATCATCCAAAAGACTTCCTTAAGTAAGATTTTCCATATTCCTTG 251
 181 AAAATGTTATCATCATCCAAAAGACTTCCTTAAGTAAGATTTTCCATATTCCTTG 240
 252 GAACATCTTCAGACTTCTTACTGTGGGCTTGTCCGAGTTCATATGCTGTTTGCTTA 311
 241 GAACATCTTCAGACTTCTTACTGTGGGCTTGTCCGAGTTCATATGCTGTTTGCTTA 300
 312 AAAAGCCTTAGGAAATAGACTTGAAGTCAACCAATATAAAAAGCTTCAGCTACAATT 371
 301 AAAAGCCTTAGGAAATAGACTTGAAGTCAACCAATATAAAAAGCTTCAGCTACAATT 360
 372 GGAGAGCTCATACACCTTCAAGAACTTAACTGAATGACAACTCACTTGGAGTCAATTAGT 431
 361 GGAGAGCTCATACACCTTCAAGAACTTAACTGAATGACAACTCACTTGGAGTCAATTAGT 420
 432 GTAGCCTTGTGTCATTCACACTCCAGAACTCACTTCGGAGTTTGGACCTCAGCAAGAAC 491
 421 GTAGCCTTGTGTCATTCACACTCCAGAACTCACTTCGGAGTTTGGACCTCAGCAAGAAC 480
 492 AAAATCAAGCACTCCCTGTGAGTTTTCAGCTCCAGAACTTAAAGATTTAAACCTT 551
 481 AAAATCAAGCACTCCCTGTGAGTTTTCAGCTCCAGAACTTAAAGATTTAAACCTT 540
 552 GACGATAATGAATTTGATTCATTTCTTTCAGATAGGACAACTAATAACCTTCGCTTT 611
 541 GACGATAATGAATTTGATTCATTTCTTTCAGATAGGACAACTAATAACCTTCGCTTT 600

QY 612 TTGTGAGCAGCTCGAAATAAGCTTCCATTTTTCCTAGTGAATTTAGAAATTTATCCCTT 671
 Db 601 TTGTGAGCAGCTCGAAATAAGCTTCCATTTTTCCTAGTGAATTTAGAAATTTATCCCTT 660
 QY 672 GAATACTTGGATCTTTTGGAAATCTTTTGAACACCAAAAGTCTTCCAGTAAATAG 731
 Db 661 GAATACTTGGATCTTTTGGAAATCTTTTGAACACCAAAAGTCTTCCAGTAAATAG 720
 QY 732 CTGCAAGCACCATTAACTTTTATTTGAAATCTTCTGCAAGAACCATTTATCATATAGG 788
 Db 721 CTGCAAGCACCATTAACTTTTATTTGAAATCTTCTGCAAGAACCATTTATCATATAGG 777

RESULT 9
 AAS56409
 ID AAS56403 standard; cDNA; 682 BP.
 XX
 AC AAS56403;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human cDNA for an ovarian cancer protein #33.
 XX
 KW Human; ss; ovarian cancer protein; cancer; tumor; ovarian cancer;
 KW endometrial cancer; cytostatic.
 XX
 OS Homo sapiens.
 XX
 FN WO200170976-A2.
 XX
 PD 27-SEP-2001.
 XX
 XX 20-MAR-2001; 2001WO-US009062.
 XX
 XX 21-MAR-2000; 2003US-0190710P.
 PR 22-JUN-2000; 2003US-0213748P.
 PR 19-DEC-2000; 2003US-0257276P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX
 PI Xu J, Pyle RA, Stolk JA;
 XX
 XX WPI; 2001-607531/69.
 DR
 XX
 XX Nucleic acids encoding 222 polypeptides associated with ovarian and
 PT endometrial cancers, useful for diagnosing, preventing and treating
 PT cancers.
 XX
 PS Claim 1; Page 126; 187pp; English.

The invention relates to human polynucleotides encoding proteins associated with ovarian and endometrial cancers. The polynucleotides and the proteins they encode may be used in the prevention, diagnosis and treatment of diseases associated with the inappropriate expression of ovarian and endometrial cancer polypeptides (OECs). For example, the polynucleotide (or an expression vector comprising the polynucleotide) and the OEC may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of OECs by expressing inactive proteins or to supplement the patient's own production of them. Additionally, the polynucleotide may be used to produce the OECs, by inserting the nucleic acids into a host cell and culturing the cell to express the protein. The polynucleotide and its complementary sequences may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acids in samples, and therefore which patients may be in need of restorative therapy. The OECs may also be used as antigens in the production of anti-OEC antibodies and in assays to identify modulators of its expression and activity. The anti-OEC antibodies and antagonists may also be used to down regulate expression and activity. The anti-OEC antibodies may also be used as diagnostic agents for detecting the presence of OEC in samples (e.g. by enzyme linked immunosorbent assay (ELISA)) and hence diagnose patients with cancers. The present sequence is a ovarian and endometrial cancer linked cDNA of

```
CC the invention
XX
SQ Sequence 682 BP; 200 A; 140 C; 131 G; 200 T; 0 U; 11 Other;

Query Match      42.9%; Score 521.6; DB 4; Length 682;
Best Local Similarity 97.5%; Pred. No. 9.9e-126;
Matches 551; Conservative 0; Mismatches 10; Indels 4; Gaps 2;

QY 277 GCGTGTGCGAGTTGATATGCGTATGCTTTGCTTTAAAGAGCCTTAGGAATTAGACTTGA 336
Db 1 GCGTGTGCGAGTTGATATGCGTATGCTTTGCTTTAAAGAGCCTTAGGAATTAGACTTGA 60

QY 337 GTCACACCATATAAAGAGCTTCCAGCTACATTTGGAGACCTTCATACACCTTCAAGAAC 396
Db 61 GTCACACCATATAAAGAGCTTCCAGCTACATTTGGAGACCTTCATACACCTTCAAGAAC 120

QY 397 TTAACCTGAATGACATCTCTGGAGTCAATTTAGTGTAGCTTGTGTCTATCTACACTCC 456
Db 121 TTAACCTGAATGACATCTCTGGAGTCAATTTAGTGTAGCTTGTGTCTATCTACACTCC 180

QY 457 AGAAGTCACTTCGGAGTTTGGACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGCGAGT 516
Db 181 AGAAGTCACTTCGGAGTTTGGACCTCAGCAAGAACAAATCAAGGCACTCCCTGTGCGAGT 240

QY 517 TTGCGAGCTCCAGGAACCTTAAGAAATTTAAACTTCAGCATTAATGAATTGATTCATTTTC 576
Db 241 TTGCGAGCTCCAGGAACCTTAAGAAATTTAAACTTCAGCATTAATGAATTGATTCATTTTC 300

QY 577 CTTGCAAGATAGACAACTTAATAAACCTTCGCTTTTGTGACGAGCTCGAAATAAGCTTC 636
Db 301 CTTGCAAGATAGACAACTTAATAAACCTTCGCTTTTGTGACGAGCTCGAAATAAGCTTC 360

QY 637 CATTTTGCCTAGTCAATTTAGAAATTTATCCCTTCAATACCTTCTTTTGGAAATA 696
Db 361 CATTTTGCCTAGTCAATTTAGAAATTTATCCCTTCAATACCTTCTTTTGGAAATA 420

QY 697 CTTTGAACCAACAAAGCTTCCAGTAAATAAGCTGCAAGCACCATAACTTTATTGG 756
Db 421 CTTTGAACCAACAAAGCTTCCAGTAAATAAGCTGCAAGCACCATAACTTTATTGG 480

QY 757 AATCTTCTGCGAGCACTTATACATAAT-AGGATTCATATGGCTCT---CATATCAT 812
Db 481 AATCTTCTGCGAGCACTTATACATAAT-AGGATTCATATGGCTCTCTTCATATCAT 540

QY 813 CCATTCATCTCTGCCAAGATTGG 837
Db 541 CCATTCATCTCTGCCAAGATTGG 565

RESULT 10
ABL68766/c
ID ABL68766 standard; DNA; 467 BP.
XX
AC ABL68766;
XX
DT 15-MAY-2002 (first entry)
XX
DE Kidney cancer related gene sequence SEQ ID NO:7103.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; anticarcinoma; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
CS Homo sapiens.
XX
PN W0200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US0-0838.
XX
PR 05-JUN-2000; 2000US-0209473P.
```

```
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, and determining a change
XX in expression of a gene of a signature gene set.
XX
XX Claim 1; SEQ ID NO 7103; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
XX neoplastic agent. The method involves exposing cells to a chemical agent
XX to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening an
XX anti-neoplastic agent, and can be used for producing a product which is
XX the data collected with respect to the anti-neoplastic agent as a result
XX of M1, and the data is sufficient to convey the chemical structure and/or
XX properties of the agent. M1 can be used in the treatment of cancer such
XX as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
XX prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
XX cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
XX cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
XX tumour
XX
```


QY 1206 AAAAAAAAAA 1215
 Db 1180 ACAAACACA 1189

RESULT 12
 ID AAH09329/c
 ID AAH09329 standard; cDNA; 584 BP.
 AC AAH09329;
 XX
 XX
 XX 26-JUN-2001 (first entry)
 XX
 XX Human cDNA clone (3'-primer) SEQ ID NO:6164.
 DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 XX Homo sapiens.
 XX
 XX EP1074617-A2.
 XX
 XX C7-FEB-2001.
 XX
 XX 28-JUL-2000; 2000EP-00116126.
 XX
 XX 29-JUL-1999; 99JP-00248036.
 XX
 XX 27-AUG-1999; 99JP-00300253.
 XX
 XX 11-JAN-2000; 2000JP-00118776.
 XX
 XX 02-MAY-2000; 2000JP-00183767.
 XX
 XX 09-JUN-2000; 2000JP-00241899.
 XX
 XX (HELI-) HELIX RES INST.
 XX
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 XX
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 XX
 XX Claim 3; SEQ ID NO 6164; 2537pp + Sequence Listing; English.
 XX
 XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
 XX
 XX Sequence 584 BP; 177 A; 132 C; 121 G; 146 T; 0 U; 8 Other;
 SQ
 Query Match 34.5%; Score 418.9; DB 4; Length 584;
 Best Local Similarity 99.5%; Pred. No. 5.8e-99;
 Matches 420; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 782 TATAGGATTCATATGGCTCTCATATCATTTCCATTCATCTCTGCCAAGATTGGATAC 841
 Db 422 TATGAGGATTCATATGGCTCTCATATCATTTCCATTCATCTCTGCCAAGATTGGATAC 363
 QY 842 CGCAAAATTTGTTGTGGAGATCTCTCTGAACTCTTTTCAATTCAGGAACTACTAC 901
 Db 362 CGCAAAATTTGTTGTGGAGATCTCTCTGAACTCTTTTCAATTCAGGAACTACTAC 303
 QY 902 CATGAATCTGCATCTCTGTTGCCACACTGTGGTCTTAGTAGATAAATTTGGTGGTACTGA 961
 Db 302 CATGAATCTGCATCTCTGTTGCCACACTGTGGTCTTAGTAGATAAATTTGGTGGTACTGA 243
 QY 962 AGCACCTATATCTCTTATTTCTGTTCTCTAGGCTGTATGTTAAATTCCTCTCATATGTT 1021
 Db 242 AGCACCTATATCTCTTATTTCTGTTCTCTAGGCTGTATGTTAAATTCCTCTCATATGTT 183
 QY 1022 AAGTAATGGTSGAGACCAAGAAAAGAAATTTCAATAACAGATCAGTTTGGGGTGCATGT 1081
 Db 182 AAGTAATGGTSGAGACCAAGAAAAGAAATTTCAATAACAGATCAGTTTGGGGTGCATGT 123
 QY 1082 ATGATTTTGCAGGCTCAAAATTTGGAGTAAGGAAAGATTTCTGTATATCTGTGGAGAGGAG 1141
 Db 122 ATGATTTTGCAGGCTCAAAATTTGGAGTAAGGAAAGATTTCTGTATATCTGTGGAGAGGAG 63
 QY 1142 GAATGTATAGTACTCTCTATTTAGATGACTCCAAAACCTTTTATTAACCAATTTAGTT 1201
 Db 62 GAATGTATAGTACTCTCTATTTAGATGACTCCAAAACCTTTTATTAACCAATTTAGTT 3
 QY 1202 TT 1203
 Db 2 TT 1

RESULT 13
 AAH15353
 ID AAH15353 standard; cDNA; 1590 BP.
 XX
 XX AC AAH15353;
 XX
 XX DT 26-JUN-2001 (first entry)
 XX
 XX DE Human cDNA sequence SEQ ID NO:13527.
 XX
 XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 XX OS Homo sapiens.
 XX
 XX PN EP1074617-A2.
 XX
 XX PD C7-FEB-2001.
 XX
 XX PF 28-JUL-2000; 2000EP-00116126.
 XX
 XX PR 29-JUL-1999; 99JP-00248036.
 XX
 XX PR 27-AUG-1999; 99JP-00300253.
 XX
 XX PR 11-JAN-2000; 2000JP-00118776.
 XX
 XX PR 02-MAY-2000; 2000JP-00183767.
 XX
 XX PR 09-JUN-2000; 2000JP-00241899.
 XX
 XX (HELI-) HELIX RES INST.
 XX
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 XX
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 XX
 XX Claim 8; SEQ ID NO 13527; 2537pp + Sequence Listing; English.

This invention describes novel isolated human polynucleotides obtained by screening for differential expression in colon cancer, breast cancer and lung cancer cell lines. The polynucleotides of the invention are represented in AAX98275-X99118 and encode polypeptides of protein families selected from 4 transmembrane segments integral membrane proteins, 7 transmembrane receptors, Apases associated with various cellular activities (AAA), eukaryotic aspartyl proteases, GATA family of transcription factors, G-protein alpha subunit, phorbolsters or diacylglycerol binding proteins, protein kinase, protein phosphatase 2C, protein tyrosine phosphatase, trypsin, wnt family of developmental signalling proteins and WW/rsp5/WWP domain containing proteins. The encoded polypeptides also have a functional domain selected from Ank repeat, basic region plus leucine zipper transcription factors, bromodomain, EF-hand, SH3 domain, WD domain/G-beta repeats, zinc finger (C2H2 type), zinc finger (CCHC class), and zinc-binding metalloprotease domain. The polynucleotides encode polypeptides with similarity to known protein families and are predicted to have similar properties. The novel polynucleotides can be used to develop products for use as therapeutic agents and in forensics, genetic analysis, mapping and diagnostic applications. In particular, the product can be used for the detection and management of cancers. They can be used for treating e.g. cervical cancers, melanomas, colorectal adenocarcinomas, Wilm's tumour, sarcomas, retinoblastoma, myosarcomas, lung carcinomas, leukemias, such as chronic myelogenous leukemia, lymphocytic leukemia, monocytic leukemia, and myeloid leukemia, and lymphomas such as histiocytic lymphoma anhydric

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2004, 21:45:13 ; Search time 3273 Seconds
(without alignments)
16089.761 Million cell updates/sec

Title: US-10-009-557-34
Perfect score: 1215
Sequence: 1 gaagaactagcatgtatgta.....ttagttaaataaaaaa 1215

Scoring table: IDENTITY_NJC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.scs.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rtd.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1215	100.0	1215	6	BD276293	BD276293 EXTRACELL
2	1215	100.0	1215	6	AX047348	AX047348 Sequence
3	1153.2	94.9	1471	6	AR339608	AR339608 Sequence
4	1153.2	94.9	1613	9	EC064133	EC064133 Homo sapi
5	1138	93.7	1445	9	AF052405	AF052405 Homo sapi
6	986.8	81.2	1950	9	BSM08175	BSM08175 Homo sapi
7	788.6	64.9	157010	9	CNS01DX7	AL139099 Human chr
8	566.6	46.8	193800	2	AC126994	AC126994 Rattus no
9	566	46.6	232458	2	AC099934	AC099934 Mus muscu
10	521.6	42.9	682	6	AX255882	AX255882 Sequence
11	443.8	36.5	467	6	AX336594	AX336594 Sequence
12	430	35.4	990	9	BC030142	BC030142 Homo sapi
13	418.8	34.5	584	6	AX871259	AX871259 Sequence
14	418.8	34.5	584	6	BD151321	BD151321 Primer fo
15	418.8	34.5	1590	6	AX878622	AX878622 Sequence
16	418.8	34.5	1590	6	BD157345	BD157345 Primer fo
17	418.8	34.5	1590	9	AK027315	AK027315 Homo sapi
18	416.2	34.3	75921	2	AC107185_3	Continuation (4 of
19	387	31.9	1681	5	BC064158	BC064158 Silurana
20	281.4	23.2	750	6	BD129164	BD129164 Human gen
21	281	23.1	732	6	BD129168	BD129168 Human gen
22	212.8	17.5	751	10	BC022648	BC022648 Mus muscu
23	188.8	15.5	770	6	BD215981	BD215981 Novel hum
24	115.4	9.5	132396	9	AL607023	AL607023 Human DNA
25	91.6	7.5	430	6	AX070036	AX070036 Sequence
26	85.2	7.0	382	6	AX070144	AX070144 Sequence
27	84.4	6.9	236640	2	AC114464	AC114464 Rattus no
28	84.4	6.9	270622	2	AC115235	AC115235 Rattus no
29	66	5.4	227056	2	AC121316	AC121316 Mus muscu
30	60.2	5.0	2950	6	AX552112	AX552112 Sequence
31	59.6	4.9	1658	9	AF359380	AF359380 Homo sapi
32	59.6	4.9	2156	9	BC003193	BC003193 Homo sapi
33	59.6	4.9	2886	9	AF332199	AF332199 Homo sapi
34	56.2	4.6	2169	6	BD251848	BD251848 RING fing
35	56.2	4.6	2169	6	AR224072	AR224072 Sequence
36	55	4.5	4941	9	BSM087337	BSM087337 Homo sapi
37	55	4.5	4984	6	AX556708	AX556708 Sequence
38	54.4	4.5	2694	6	BD232164	BD232164 A novel m
39	53.8	4.4	11990	1	AE010918	AE010918 Methanosa
40	53.8	4.4	164026	2	AC092762	AC092762 Pan trogl
41	53.8	4.4	173456	2	AC092762	AC092762 Pan trogl
42	53.4	4.4	1141	6	AX083744	AX083744 Sequence
43	53.4	4.4	128833	2	AC135793	AC135793 Oryza sat
44	52.4	4.3	2620	6	AX877169	AX877169 Sequence
45	52.4	4.3	2620	6	BD156521	BD156521 Primer fo

ALIGNMENTS

RESULT 1
BD276293
LOCUS
DEFINITION
EXTRACELLULAR MATRIX AND ADHESION-ASSOCIATED PROTEINS.
ACCESSION
BD276293
VERSION
BD276293.1
KEYWORDS
JP 2002543785-A/9.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1215)
AUTHORS
Azimzai, Y., Bandman, O., Tang, T.Y., Lai, P., Henry, Yue, Baughn, M.R.,
Lu, D.A.M. and Hillman, J.L.
TITLE
EXTRACELLULAR MATRIX AND ADHESION-ASSOCIATED PROTEINS

Db 61 TTCCAGGCCATTTCCAGCAGTTTAAAGGTTTCTTTTCAGCTATGAGACTGGCTCATGAG 120
QY 121 GCTGTAATGTTGATACACGATTTTCAACGCTCACACAGTGAAGACTTCAGAAATTCGAAA 180
Db 121 GCTGTAATGTTGATACACGATTTTCAACGCTCACACAGTGAAGACTTCAGAAATTCGAAA 180
QY 181 ACTTTAAACCTAAATGTTTATCACATCCAAAAGAGACTATCTCTAAGTAAGAAATTTTC 240
Db 181 ACTTTAAACCTAAATGTTTATCACATCCAAAAGAGACTATCTCTAAGTAAGAAATTTTC 240
QY 241 CATATTTCTTTGGAACTCTTACAGACTTTTCTTACTGTGGCTTTGTCGAGTTGATATGGTA 300
Db 241 CATATTTCTTTGGAACTCTTACAGACTTTTCTTACTGTGGCTTTGTCGAGTTGATATGGTA 300
QY 301 TCGTTTGTCTTAAAGAGCTTTAGAAATAGACTTTGAGTCAACCCATATATAAAGAGCTTC 360
Db 301 TCGTTTGTCTTAAAGAGCTTTAGAAATAGACTTTGAGTCAACCCATATATAAAGAGCTTC 360
QY 361 CAGCTACAAATTCGAGACCTCATACACCTTCAAGAACTTAACTGAATGACAACTCACTGG 420
Db 361 CAGCTACAAATTCGAGACCTCATACACCTTCAAGAACTTAACTGAATGACAACTCACTGG 420
QY 421 AGTCATTTAGTGTAGCTTTGTGTCTTACTTACACTCCAGAGTCACTTCGAGTTTGGACC 480
Db 421 AGTCATTTAGTGTAGCTTTGTGTCTTACTTACACTCCAGAGTCACTTCGAGTTTGGACC 480
QY 481 TCAGCAAGAACAAATCAAGGACCTCCCTGTGAGTTTTCGAGCTTCAGGAACTTAAAGA 540
Db 481 TCAGCAAGAACAAATCAAGGACCTCCCTGTGAGTTTTCGAGCTTCAGGAACTTAAAGA 540
QY 541 ATTTAAACCTCAGCAATATGAATGATCAATTTCTTTCGAGATGAGCACTAAATAA 600
Db 541 ATTTAAACCTCAGCAATATGAATGATCAATTTCTTTCGAGATGAGCACTAAATAA 600
QY 601 ACTTCGCTTTTGTGAGCAGCTCGAATTAAGCTTCCATTTTTCGCTAGTGAATTTAGAA 660
Db 601 ACTTCGCTTTTGTGAGCAGCTCGAATTAAGCTTCCATTTTTCGCTAGTGAATTTAGAA 660
QY 661 ATTTATCCCTTGAATGATGATTTTGGAAATCTTTTGAAGAACTTTTGAACCAACCAAGTCCCTTC 720
Db 661 ATTTATCCCTTGAATGATGATTTTGGAAATCTTTTGAAGAACTTTTGAACCAACCAAGTCCCTTC 720
QY 721 CAGTAATAAGAGCTGCAAGCACCATTAACTTTATGGAATCTTCTGCAAGAACCATATTAC 780
Db 721 CAGTAATAAGAGCTGCAAGCACCATTAACTTTATGGAATCTTCTGCAAGAACCATATTAC 780
QY 781 ATAACTAGGAATTCATATGCTCTATATCATTCATTCATCTCTGCAAGAAATTTGGATA 840
Db 781 ATAACTAGGAATTCATATGCTCTATATCATTCATTCATCTCTGCAAGAAATTTGGATA 840
QY 841 CCGCAAAAATTTGTTTGTGGAAGATCTGTCTGCAACTCTTTTCAATTCAGGAATCTACTA 900
Db 841 CCGCAAAAATTTGTTTGTGGAAGATCTGTCTGCAACTCTTTTCAATTCAGGAATCTACTA 900
QY 901 CCATGAATCTGCAATCTCTGTCGCCACACCTGTGCTTAGTAGATAAATTTGGTGGTACTG 960
Db 901 CCATGAATCTGCAATCTCTGTCGCCACACCTGTGCTTAGTAGATAAATTTGGTGGTACTG 960
QY 961 AAGCACTTATATCTTTATTTCTTCTCTAGGCTGTATGTTAACTTCTCTGATATGT 1020
Db 961 AAGCACTTATATCTTTATTTCTTCTCTAGGCTGTATGTTAACTTCTCTGATATGT 1020
QY 1021 TAAAGTAATGGGTGAGACCAAGAAATTTCAATAACAGATCAGTTTGGGGTGCATG 1080
Db 1021 TAAAGTAATGGGTGAGACCAAGAAATTTCAATAACAGATCAGTTTGGGGTGCATG 1080
QY 1081 TATGATTTTCAGGCTCAAAATTTGGAGTAAGGAAGATTTCTGTATATCTTGTCTGGAGAGGA 1140
Db 1081 TATGATTTTCAGGCTCAAAATTTGGAGTAAGGAAGATTTCTGTATATCTTGTCTGGAGAGGA 1140
QY 1141 GGAATGTATGATGATCTATTTAGATGACCTCCAAAGCTTTTATTAACCAATTTTGT 1200

Db 1141 GGAATGTATGATGATCTATTTAGATGACTCCAAAAGCTTTTATTAACCAATTTTGT 1200
QY 1201 TTTAAAAAAGAAAAA 1215
Db 1201 TTTAAAAAAGAAAAA 1215
RESULT 3
AR339608 1471 bp DNA linear PAT 17-AUG-2003
LOCUS
DEFINITION Sequence 1099 from patent US 6569662.
ACCESSION AR339608
VERSION AR339608.1 GI:33726465
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1471)
AUTHORS Tang, Y.T., Zhou, P. and Drmanac, R.T.
TITLE Nucleic acids and polypeptides
JOURNAL Patent: US 6569662-A 1099 27-MAY-2003;
FEATURES Location/Qualifiers
source 1..1471
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 94.9%; Score 1153.2; DB 6; Length 1471;
Best Local Similarity 99.7%; Pred. No. 1-2e-247;
Matches 1155; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 58 TTTATTCAGGCCATTTCCAGCAGTTTAAAGGTTTCTTTTCAGCTATGAGACTGGCTCAT 117
Db 314 TAACTAAGGCCATTTCCAGCAGTTTAAAGGTTTCTTTTCAGCTATGAGACTGGCTCAT 373
QY 118 GAGGCTGTAAATGTTGATACACAGTTTCAACGCTTCACCCAGTGAGACTTCAGAAATTC 177
Db 374 GAGGCTGTAAATGTTGATACACAGTTTCAACGCTTCACCCAGTGAGACTTCAGAAATTC 433
QY 178 AAAAATTTAAACTAAATGGTTATCAGATCCAAAAAGACTATCTCTAAGTAAGAAAT 237
Db 434 AAAAATTTAAACTAAATGGTTATCAGATCCAAAAAGACTATCTCTAAGTAAGAAAT 493
QY 238 TTCCATATTTCTTGGAGACTTCAGACTTCTTACTGTGGCTTTCAGGATGATATGC 297
Db 494 TTCCATATTTCTTGGAGACTTCAGACTTCTTACTGTGGCTTTCAGGATGATATGC 553
QY 298 GTATGCTTTGCTTAAAGAGCTTTAGGAATTAGACTTTGAGTCAACACCATATAAAAAAGC 357
Db 554 GTATGCTTTGCTTAAAGAGCTTTAGGAATTAGACTTTGAGTCAACACCATATAAAAAAGC 613
QY 358 TTCCAGCTTAAATTTGAGACTTCATACACCTTCAAGAACTTAACTTAACTGAAATGCAATCACT 417
Db 614 TTCCAGCTTAAATTTGAGACTTCATACACCTTCAAGAACTTAACTTAACTGAAATGCAATCACT 673
QY 418 TGGAGTCAATTTAGTGTAGCTTGTCTTACTTACACTCCAGAGTCACTTCGAGTTGG 477
Db 674 TGGAGTCAATTTAGTGTAGCTTGTCTTACTTACACTCCAGAGTCACTTCGAGTTGG 733
QY 478 ACCTCAGCAAGAACAAAAATCAAGGACATCCCTGTGCTAGTTTTCAGGCTCCAGGAATCTA 537
Db 734 ACCTCAGCAAGAACAAAAATCAAGGACATCCCTGTGCTAGTTTTCAGGCTCCAGGAATCTA 793
QY 538 AGAATTTAAACTTGAAGATTAATGAATTTCAATTTCCCTTGCAGATAGGACAACTAA 597
Db 794 AGAATTTAAACTTGAAGATTAATGAATTTCAATTTCCCTTGCAGATAGGACAACTAA 853
QY 598 TAAACCTTTCGGTTTGTGTCAGCAGCTCGAATTAAGCTTCCATTTTTCCTAGTGAATTA 657
Db 854 TAAACCTTTCGGTTTGTGTCAGCAGCTCGAATTAAGCTTCCATTTTTCCTAGTGAATTA 913
QY 658 GAAATTTATCCCTTGAATTAATTTGGATCTTTTGGAAATACTTTTGAACCAACCAATTTTGT 717

```

Db      914  GAAATTTATCCCTGATACCTGGATCTTTTGGAAATACCTTTGAACACCAAAAAGTCC 973
QY      718  TTCAGTAATAAGCTGCAAGACCAATTAATTTATTTGGAAATCTTCTGCGAAGAACCATAT 777
Db      974  TTCAGTAATAAGCTGCAAGACCAATTAATTTATTTGGAAATCTTCTGCGAAGAACCATAT 1033
QY      778  TACATAATAGATTCATATAGCTGCTCATATCATTCATTCATCCATCTGCGAAGATTTGG 837
Db      1034  TACATAATAGATTCATATAGCTGCTCATATCATTCATTCATCCATCTGCGAAGATTTGG 1093
QY      838  ATACCGCAAAAATTTGTTTGGGAAGATCTCTCTGAATCTTCTTCAATCAAGGAACCTA 897
Db      1094  ATACCGCAAAAATTTGTTTGGGAAGATCTCTCTGAATCTTCTTCAATCAAGGAACCTA 1153
QY      898  CTACCATGATCTGCATCTGTTGCCACACATGCTGCTAGTAGATAATTTGGTGGTGA 957
Db      1154  CTACCATGATCTGCATCTGTTGCCACACATGCTGCTAGTAGATAATTTGGTGGTGA 1213
QY      958  CTGAAGCACCTATTATCTCTTATTTCTGTTCTCTAGGCTGTTAGTTAATTTCTCTGATA 1017
Db      1214  CTGAAGCACCTATTATCTCTTATTTCTGTTCTCTAGGCTGTTAGTTAATTTCTCTGATA 1273
QY      1018  TGTAAAGTAATGGTGGAGACCGAAGAAATTTCAATAACAGATCAGTTTGGGTGC 1077
Db      1274  TGTAAAGTAATGGTGGAGACCGAAGAAATTTCAATAACAGATCAGTTTGGGTGC 1333
QY      1078  ATCTATGATTTTCAGGCTCAATTTGGAGTAGAGTAATTTCTGATATCTGCTGAGA 1137
Db      1334  ATCTATGATTTTCAGGCTCAATTTGGAGTAGAGTAATTTCTGATATCTGCTGAGA 1393
QY      1138  GGAGGAATGATAGTATCTATCTATTTAGATGACTCCAAAATTTTATTAACCAATTTT 1197
Db      1394  GGAGGAATGATAGTATCTATCTATTTAGATGACTCCAAAATTTTATTAACCAATTTT 1453
QY      1198  AGTTTAAAAAAAATAA 1215
Db      1454  AGTTTAAAAAAAATAA 1471

RESULT 4
LOCUS   BC064133
DEFINITION Homo sapiens cDNA clone MGC:74569 IMAGE:6501107, complete cds.
ACCESSION BC064133
VERSION   BC064133.1 GI:39794694
KEYWORDS MGC.
SOURCE   Homo sapiens
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1613)
Straussberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altshul,S.P., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Udwin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Huiyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallos,D.E.,
Schnerker,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932

```

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 1613)

Straussberg,R.
Direct Submission
Submitted (11-DEC-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Varon Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lilisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smallos, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC P-rate: 141 Row: 1 Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 22748718.

FEATURES

source

1. 1613
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:74569 IMAGE:6501107"
/tissue_type="Uterus, leiomyosarcoma"
/clone_lib="NIH_MGC_71"
/lab_host="DH10B"
/notes="Vector: pOMV-SPORT6"
70..264
/codon_start=1
/product="Unknown (protein for MGC:74569)"
/protein_id="AAH64133.1"
/db_xref="GI:39794694"
/translation="MKLHCEVEVISRHLPALGLNRNKGVRVAVLSLCQTSRSPQPPVR
AFLLISTLXDKRGTRYHSG"

CDS

ORIGIN

Query Match 94.9%; Score 1153.2; DB 9; Length 1613;
Best Local Similarity 99.7%; Pred. No. 1.2e-247;
Matches 1155; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 58 TTATTCAGGCCCAATTCAGCAGTTTAAAGGTTTCCCTTCAGCTATGAGACTGGCTCAT 117
Db 452 TAAGTAAGGCCAATTCAGCAGTTTAAAGGTTTCCCTTCAGCTATGAGACTGGCTCAT 511
QY 118 GAGGCTGTAATGTTGATACACCGTTTCAACGCTCACACAGTGAAGACTTCAGAAATTG 177
Db 512 GAGGCTGTAATGTTGATACACCGTTTCAACGCTCACACAGTGAAGACTTCAGAAATTG 571
QY 178 AAAACTTTAAACTAAATGTTTATCACATCCAAAAGACTATCTCTTAAAGTAAGAAATT 237
Db 572 AAAACTTTAAACTAAATGTTTATCACATCCAAAAGACTATCTCTTAAAGTAAGAAATT 631
QY 238 TTCCATATTCCTTGGAAACATCTTCAGACTTCTTACTGTGGGCTTGTCCGAGTTGATATGC 297
Db 632 TTCCATATTCCTTGGAAACATCTTCAGACTTCTTACTGTGGGCTTGTCCGAGTTGATATGC 691
QY 298 GTATGCTTGTCTTAAAGCCTTAGAAATTAGACTTGTAGTCAACACCATATAAAAAGC 357
Db 692 GTATGCTTGTCTTAAAGCCTTAGAAATTAGACTTGTAGTCAACACCATATAAAAAGC 751

TITLE
JOURNAL
PUBMED

QY	358	TTCCAGCTACAAATTAGTGAGACCTCATACACCTTCAAGAACTTAACTGAATGCAATCACT	4117
Db	752	TTCCAGCTACAAATTAGTGAGACCTCATACACCTTCAAGAACTTAACTGAATGCAATCACT	8111
QY	418	TGGAGTCAATTTAGTCTAGCCTTGCTCAATTCATACACTCCAGAACTCACTTCGGAGTTGG	4777
Db	812	TGGAGTCAATTTAGTGTAGCCTTGCTCAATTCATACACTCCAGAACTCACTTCGGAGTTGG	871
QY	478	ACCTCAGCAAGAAACAAAATCAAGGCACCTCCCTGTGTCAGTTTGTGCAGCTCCAGGAACCTTA	537
Db	872	ACCTCAGCAAGAAACAAAATCAAGGCACCTCCCTGTGTCAGTTTGTGCAGCTCCAGGAACCTTA	931
QY	538	AGAAATTTAAACCTTGACGATTAATGAATGATCAATTTCTTGCAGATAGGAACAATAA	597
Db	932	AGAAATTTAAACCTTGACGATTAATGAATGATCAATTTCTTGCAGATAGGAACAATAA	991
QY	598	TAAACCTTCGCTTTTGTGCAGCAGCTCGAAATAAGCTTCCAAATTTTGTGCTAGTGAATTTA	657
Db	992	TAAACCTTCGCTTTTGTGCAGCAGCTCGAAATAAGCTTCCAAATTTTGTGCTAGTGAATTTA	1051
QY	658	GAATTTATCCCTTGAAATPACTTGGATCTTTTGTGAAATPACTTTTGAACAACCAAAAGTCC	717
Db	1052	GAATTTATCCCTTGAAATPACTTGGATCTTTTGTGAAATPACTTTTGAACAACCAAAAGTCC	1111
QY	718	TTCCAGTAATAAAGCTCAAGCAACCATTAACCTTTTATTGGATCTCTCTGACGACCAATAT	777
Db	1112	TTCCAGTAATAAAGCTCAAGCAACCATTAACCTTTTATTGGATCTCTCTGACGACCAATAT	1171
QY	778	TACATAATAGGAATCCATATGGCTCTCATATCATATCCAAATCCATCTCTGCCAAGATTTGG	837
Db	1172	TACATAATAGGAATCCATATGGCTCTCATATCATATCCAAATCCATCTCTGCCAAGATTTGG	1231
QY	838	ATACCGCAAAATTTGTGTTTGTGCGAAGATCTGCTCGAACTCTTTCAATTCAGGAACCTA	897
Db	1232	ATACCGCAAAATTTGTGTTTGTGCGAAGATCTGCTCGAACTCTTTCAATTCAGGAACCTA	1291
QY	898	CTACCATGAATCTGCATTTCTGTTCGCCACACCTGTGTGCTTAGTAGATAATTTGGGTGGTA	957
Db	1292	CTACCATGAATCTGCATTTCTGTTCGCCACACCTGTGTGCTTAGTAGATAATTTGGGTGGTA	1351
QY	958	CTGAGCAGCACTTAATCTCTATTCTCTGTTCTTAGCTGTATGTTAATTCCTCTGATA	1017
Db	1352	CTGAGCAGCACTTAATCTCTATTCTCTGTTCTTAGCTGTATGTTAATTCCTCTGATA	1411
QY	1018	TGTTAAAGTAATGGGTGAGACCGAGAAAAGAAAATTCATAACACAGATCAGTTTGGGTGC	1077
Db	1412	TGTTAAAGTAATGGGTGAGACCGAGAAAAGAAAATTCATAACACAGATCAGTTTGGGTGC	1471
QY	1078	ATGTATGAATTTGACAGGTCAATTTGAGTAAGGAAGATTCTCTGTATATCTCTGTGAGA	1137
Db	1472	ATGTATGAATTTGACAGGTCAATTTGAGTAAGGAAGATTCTCTGTATATCTCTGTGAGA	1531
QY	1138	GGAGGAATCTGTATAGTTACTCATTTAGATGACTCCAAAACCTTTTATTAACCAATTTT	1197
Db	1532	GGAGGAATCTGTATAGTTACTCATTTAGATGACTCCAAAACCTTTTATTAACCAATTTT	1591
QY	1198	AGTTTTTAAAAAATAAAAA 1215	
Db	1592	AGTTTTTAAAAAATAAAAA 1609	
RESULT 5	AY052405		
LOCUS		1445 bp	linear
DEFINITION	Homo sapiens 4-LBB-mediated signaling molecule mRNA, complete cds.		
ACCESSION	AY052405		
VERSION	AY052405.1		
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 1445)		

AUTHORS	Kwon,B.S. and Jang,I.K.
TITLE	Characterization of LRR-1, a new leucine-rich repeat-containing protein involved in 4-1BB-mediated signal transduction
JOURNAL	Mol. Cells (2001) In press
REFERENCE	2 (bases 1 to 1445)
AUTHORS	Kwon,B.S. and Jang,I.K.
TITLE	Direct Submission
JOURNAL	Submitted (26-AUG-2001) Ophthalmology, Louisiana State University, 2020 Gravier Street, New Orleans, LA 70112, USA
FEATURES	Location/Qualifiers
source	1..1445
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
CDS	11..1255
	/note="4-1BB1rr; LRR-1 signaling protein"
	/codon_start=1
	/product="4-1BB-mediated signaling molecule"
	/protein_id="AAU1430.1"
	/db_xref="GI:16024938"
	/translation="MKLHCVEVVISRHLPALGNRGKVVSRVLSLQQTSRSQPPYRVC AFLLIQLDKRGTRVELRENIIQQFTFKDEGATVRLKEPPVDVILGLSKAISSLGK FLAMSLAHRGQNVDPVSTLTTPKTSSEFENKTKWITSKQYPLSKNPPYSYELHLC TSYCGLRVDMRLCLSLKRLDLSHNHKKLPATIGDLIEQLNLNDHLSFSAVH LCHSTQKSRGSLGDKSKKTKALPVQFCQQLNKLKDDNELIQFPCKGQLNLIRH LSAARMLPLPEFENFLSLEYLDLFGNTPEQKVLFPVKLOAPLTLLSSARTLIRH RIYPIGHIIPIPHLCQDLLEAKI CVCGRFCLNSIFQGTTMNLHSAVETVVLVDNLGGT EAPIISVFCSGIYGVNNSDPMKX"

ORIGIN	Query Match: Best Local Similarity Matches 1152; Conservative	93.7%; 99.5%; 0;	Score 1138; Pred. No. 3e-244; 0;	DB 9; Length 1445; 5;	Indels 1;	Gaps 1;				
QY	58	TTATT	CAGGCCA	TTCCAGCAGT	TTAAAGG	TTTCCTTTT	CAGCTAT	CAGACT	TGCTCTATA	117
Db	285	TAA	GTAAAGG	CCATTTCCAGCAGT	TTAAAGG	TTTCCTTTT	CAGCTAT	CAGACT	TGCTCTATA	344
QY	118	GAG	CGCTGTA	TGTTGATACAC	CAGGTTTCA	ACGCTCACAC	CAGTGAAG	CACTTC	CAAACTTG	177
Db	345	GAG	CGCTGTA	TGTTGATACAC	CAGGTTTCA	ACGCTCACAC	CAGTGAAG	CACTTC	CAAACTTG	404
QY	178	AAA	ACTTTTAA	ACTAAATGG	TATACAT	CCAAAAG	ACTATCC	CTTAA	GTAAGAATT	237
Db	405	AAA	ACTTTTAA	ACTAAATGG	TATACAT	CCAAAAG	ACTATCC	CTTAA	GTAAGAATT	464
QY	238	TT	CCATATTC	CTTGGAA	CACTTTT	CAGACTTCTT	ACTGTGG	CGCTTGT	CCGAGTTGATATGC	297
Db	465	TTC	CATATTC	CTTGGAA	CACTTTT	CAGACTTCTT	ACTGTGG	CGCTTGT	CCGAGTTGATATGC	524
QY	298	GTA	TGCTTTTG	CTTTAAA	AGGCTTTAG	GAAATTAG	ACTTTG	ACTCC	CAACCATATA	357
Db	525	GTA	TGCTTTTG	CTTTAAA	AGGCTTTAG	GAAATTAG	ACTTTG	ACTCC	CAACCATATA	584
QY	358	TT	CCAGCTACA	TTGGAG	CACTTAT	CACACTT	CAAGAA	CTTAA	CCCTGAGATGACATCACT	417
Db	585	TT	CCAGCTACA	TTGGAG	CACTTAT	CACACTT	CAAGAA	CTTAA	CCCTGAGATGACATCACT	644
QY	418	TG	GAGTCATT	TAGTGTAG	GCTTGTGTC	ATTCTT	ACACTCC	AGAGTCA	CTTCGAGATTGG	477
Db	645	TG	GAGTCATT	TAGTGTAG	GCTTGTGTC	ATTCTT	ACACTCC	AGAGTCA	CTTCGAGATTGG	704
QY	478	ACT	CAGCAAGA	CAAAATCA	AGGCACTCC	CTCTGTC	AGTTTGG	CCAGCTCC	AGAACTTA	537
Db	705	ACT	CAGCAAGA	CAAAATCA	AGGCACTCC	CTCTGTC	AGTTTGG	CCAGCTCC	AGAACTTA	764
QY	538	AGA	ATTTAA	ACTTGAC	GAATA	TGAATTC	ATTCAT	CAATTTCC	TTCGAGATAGGCACTAA	597
Db	765	AGA	ATTTAA	ACTTGAC	GAATA	TGAATTC	ATTCAT	CAATTTCC	TTCGAGATAGGCACTAA	824
QY	598	TAA	ACTTTG	CTTTTGT	CAGAGCTCG	AAATAAG	CTTCC	ATTTTTC	TGCTAGTGAATTTA	657
Db	825	TAA	ACTTTG	CTTTTGT	CAGAGCTCG	AAATAAG	CTTCC	ATTTTTC	TGCTAGTGAATTTA	684


```

Db      1434 AAATTGCAAAATAGGCCCATCTAGATCCCGTACCAAGAGTTCATATGGCTCTCATATC 1493
QY      810 ATTCCATTCCATCTCTGCCAAGATTTCGATACCGCAAAATTTGTGTTTGGAGATTC 869
Db      1494 ATTCCATTCCATCTCTGCCAAGATTTCGATACCGCAAAATTTGTGTTTGGAGATTC 1553
QY      870 TGTCTGAATCTTTCAATCAAGGAACATACCAATGATTCGATTCCTGTTGCCCACT 929
Db      1554 TGTCTGAATCTTTCAATCAAGGAACATACCAATGATTCGATTCCTGTTGCCCACT 1613
QY      930 GTGGCTCTAGTAGATAATTTGGTGGTACTGAGACCACTTAATCTCTTATTTCTGTCT 989
Db      1614 GTGGCTCTAGTAGATAATTTGGTGGTACTGAGACCACTTAATCTCTTATTTCTGTCT 1673
QY      990 CTAGGCTGTATCTTAATCTCTGATATCTTAAAGTAAGTGGTGGAGACCAAGAAAGAA 1049
Db      1674 CTAGGCTGTATCTTAATCTCTGATATCTTAAAGTAAGTGGTGGAGACCAAGAAAGAA 1733
QY      1050 ATTCAATAACAGATCAGTTTGGGTGCATGTATGATTTTGCAGCGTCAAAATTCGAGTAA 1109
Db      1734 ATTCAATAACAGATCAGTTTGGGTGCATGTATGATTTTGCAGCGTCAAAATTCGAGTAA 1793
QY      1110 GGGAGATTTCTGTATCTTCTGATCTGCTGGAGAGAGAACTGTATAGTACTCAATTTAGTGA 1169
Db      1794 GGGAGATTTCTGTATCTTCTGATCTGCTGGAGAGAGAACTGTATAGTACTCAATTTAGTGA 1853
QY      1170 CTCCTAAATCTTTATTAACCAATTTTGTAGTTTAAATAAAAAAAA 1215
Db      1854 CTCCTAAATCTTTATTAACCAATTTTGTAGTTTAAATAAAAAAAA 1899

RESULT 7
CNS01DX7/c
LOCUS      157010 bp DNA linear PRI 31-AUG-2001
DEFINITION Human chromosome 14 DNA sequence BAC R-649E7 of library RPCI-11
ACCESSION AL139099
VERSION    AL139099.3 GI:15424523
KEYWORDS  HTG; HGS; ACTIVEFIN.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 157010)
AUTHORS   Heilig,R., Petit,J.L., Vico,V., Dasiwa,C., Robert,C., Wincker,P.,
           Brothier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
           Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
           Gyapay,G., Saurin,W. and Weissenbach,J.
           Sequencing of the human chromosome 14
           Unpublished
           2 (bases 1 to 157010)
           Genoscope.
           Direct Submission
           Submitted (31-AUG-2001) Genoscope - Centre National de Sequencage :
           BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr
           - Web : www.genoscope.cns.fr;
           On Sep 3, 2001 this sequence version replaced gi:8217934.
           -----
           Center : Genoscope / Centre National de Sequencage
           Center code: GS
           Web site: http://www.genoscope.cns.fr/
           Contact: Seqref@genoscope.cns.fr
           -----
           The following BAC sequence is oriented from the T7 to the SP6 end.
           Upstream BAC (overlapping the T7 end) : R-831F12 (AC=AL591767)
           Downstream BAC (overlapping the SP6 end) : R-346L24
           -----
           Summary Statistics
           Assembly program: Phrap; version 2.0
           Quality coverage: 7.30x in Q20 bases; sum-of-contigs
           -----
           Overall quality chart :

```

```

Range      : bases
0          :
1 - 9      : 8
10 - 19   : 92
20 - 29   : 225
30 - 39   : 870
40 - 49   : 3443
50 - 59   : 7315
60 - 69   : 11031
70 - 79   : 25134
80 - 89   : 50478
90 - 99   : 58414

```

Percentage of bases with a quality value >= 40 : 99 %.

FEATURES
source

```

1. 157010
   location/Qualifiers
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="14"
   /clone_lib="R-649E7"
   /clone="R-649E7"
   /clone_lib="RPCI-11"
   31812..32022
   /note="matching EMBL:Z52356"
   RDB:RH31309
   dbSTS:STS15763
   Identified using the e-PCR software (G. Schuler)"
   /note="matching EMBL:N22189"
   RDB:RH69276
   dbSTS:STS49154
   Identified using the e-PCR software (G. Schuler)"
   41035..41109
   /note="matching EMBL:G31448"
   RDB:RH41760
   dbSTS:STS35761
   Identified using the e-PCR software (G. Schuler)"
   43622..43745
   /note="matching EMBL:M15661"
   RDB:RH71370
   dbSTS:STS51135
   Identified using the e-PCR software (G. Schuler)"
   43088..48262
   /note="matching EMBL:AA424400"
   RDB:RH103208
   dbSTS:STS70744
   Identified using the e-PCR software (G. Schuler)"
   64270..64448
   /note="matching EMBL:W84670"
   RDB:RH82026
   dbSTS:STS58378
   Identified using the e-PCR software (G. Schuler)"

```

ORIGIN

```

Query Match      64.9%; Score 788.6; DB 9; Length 157010;
Best Local Similarity 98.3%; Pred. No. 2.5e-166;
Matches 797; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      1 GAAGACATGACATGATGATTTATCTCCAGTGAATTTATAATTTTACAACTTTTATTA 60
Db      55298 GAAGAACATGACATGATGATTTATCTCCAGTGAATTTATAATTTTACAACTTTTATTA 55239
QY      61 TTCCAGGCCATTTCCAGCAGTTTAAAGGTTTCCCTTCAGCTATGAGACTGGCTCATAGAG 120
Db      55238 TTCCAGGCCATTTCCAGCAGTTTAAAGGTTTCCCTTCAGCTATGAGACTGGCTCATAGAG 55179
QY      121 GCTGTAATGTTGATACACACCGATTTCAACGCTCACACCACTGAGACTTCAGAAATTTGAAA 180
Db      55178 GCTGTAATGTTGATACACCGATTTCAACGCTCACACCACTGAGACTTCAGAAATTTGAAA 55119
QY      181 ACTTTAAACTAAATAGTGTATCATCATCCAAAAAGACATATCTCTAAGTAAGATTTTC 240
Db      55118 ACTTTAAACTAAATAGTGTATCATCATCCAAAAAGACATATCTCTAAGTAAGATTTTC 55059

```

QY 241 CATATTCCTTGGAAATCTTTCAGACTCTTACTGTGGGCTTGTCCAGCTTATATGCGTA 300
 |||||
 Db 55058 CATATTCCTTGGAAATCTTTCAGACTCTTACTGTGGGCTTGTCCAGCTTATATGCGTA 54999
 |||||
 QY 301 TGCCTTCTTAAAGGCTTTAGGAAATAGACTTGGTCCAAACCATATATAAAGCTTC 360
 |||||
 Db 54998 TGCCTTCTTAAAGGCTTTAGGAAATAGACTTGGTCCAAACCATATATAAAGCTTC 54939
 |||||
 QY 361 CAGTACAAATGGAGACTCATACACCTTCAAGAACTTAACCTGAATGAACAACTTGG 420
 |||||
 Db 54938 CAGTACAAATGGAGACTCATACACCTTCAAGAACTTAACCTGAATGAACAACTTGG 54879
 |||||
 QY 421 AGTCATTTAGTGTAGCTTGTGTCTCATCTACACCTCCAGAACTCACTTCGGAGTTTGACC 480
 |||||
 Db 54878 AGTCATTTAGTGTAGCTTGTGTCTCATCTACACCTCCAGAACTCACTTCGGAGTTTGACC 54819
 |||||
 QY 481 TCAGCAAGAAACAAATCAAGGCACTCCCTGTGCAAGTTTTCAGCTCCAGGAACCTTAAGA 540
 |||||
 Db 54818 TCAGCAAGAAACAAATCAAGGCACTCCCTGTGCAAGTTTTCAGCTCCAGGAACCTTAAGA 54759
 |||||
 QY 541 ATTATTAACCTTGCAGTAATGAATGATTCATCTCCAGAGATAGGACAACTAATAA 600
 |||||
 Db 54758 ATTATTAACCTTGCAGTAATGAATGATTCATCTCCAGAGATAGGACAACTAATAA 54699
 |||||
 QY 601 ACCTTCGCTTTTGTGTCAGAGCTCGAATAAGCTTCCATTTTTCCTAGTGAATTTAGAA 660
 |||||
 Db 54698 ACCTTCGCTTTTGTGTCAGAGCTCGAATAAGCTTCCATTTTTCCTAGTGAATTTAGAA 54639
 |||||
 QY 661 ATTATCCCTTGAATCTTGGATCTTTTTCGAAATACCTTTTGAACACCAAAAGTCTTTC 720
 |||||
 Db 54638 ATTATCCCTTGAATCTTGGATCTTTTTCGAAATACCTTTTGAACACCAAAAGTCTTTC 54579
 |||||
 QY 721 CAGTAATAAGCTGCAAGCACTTATCTTATGGAATCTTCTGCAAGCAACCAATATAC 780
 |||||
 Db 54578 CAGTAATAAGCTGCAAGCACTTATCTTATGGAATCTTCTGCAAGCAACCAATATAC 54519
 |||||
 QY 781 ATAATAGGATTCATATGGCTCTCATATCAT 811
 |||||
 Db 54518 ATAATAGGATTCATATGGCTCTCATATCAT 54488
 |||||

RESULT 8

AC126994
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-295C20, *** SEQUENCING IN PROGRESS ***
 AC126994
 VERSION AC126994.3 GI:25139354
 KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 REFERENCE 1 (bases 1 to 193800)
 Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, J., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Chen, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,

Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenschwa, L., Loulseghe, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindart, M., Mahmoud, M., Malloy, K., Mangum, A., Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mathiney, J., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Morgan, M., Morris, K., Morris, S., Munday, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Morris, S., Nwankweli, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfamko, C., Piopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, M., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Snajs, D., Steed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseña, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Wooley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G., and Gibbs, R. A.
 Direct Submission
 Unpublished
 2 (bases 1 to 193800)
 Morley, K. C.
 Direct Submission
 Submitted (12-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 193800)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Nov 20, 2002 this sequence version replaced gi:22855985.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: G20F
 Center clone name: CH230-235C20
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 178664 bases at least Q40
 Consensus quality: 181447 bases at least Q30
 Consensus quality: 182574 bases at least Q20
 Estimated insert size: 188087; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.bgsc.bcm.tmc.edu/docs/Genbank_graft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces
 is believed to be correct as given, however the sizes
 of the gaps between them are based on estimates that have
 been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 193800: contig of 193800 bp in length.

FEATURES

Source

Location/Qualifiers
 1..193800
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-295C20"
 1..2415
 /note="wgs_contig"
 6416..9345
 /note="wgs_contig"
 188582..190216
 /note="wgs_contig"
 191015..193800
 /note="wgs_contig"

ORIGIN

Query Match 46.8%; Score 568.6; DB 2; Length 193800;
 Best Local Similarity 83.3%; Pred. No. 4.1e-117;
 Matches 659; Conservative 0; Mismatches 129; Indels 3; Gaps 1;
 1 GAAGAACTAGCATGTATGTATATCTCCAGTGGAAATTAATTTACAA---CTTTTAT 57
 46535 GAAGAACTAGCAAGCGTGAAGTAATTAAGGTGGACCTATAATTTGTATTTT 46594
 58 TTATTCAGGCCATTTCCAGCAGTTTAAAGGTTCTTTTCAGCTATGACACTGGCTCATA 117
 46595 TTATTCAGGCCATTTCCGCAATTTAAAGACATCTTTTCAGCTATGACACTGGCTCATA 46654
 118 GAGGCTGTATGTTGATACACAGTTTCAACCTCACACGAGTGAAGCTTCAGAAATTTG 177
 46655 GAGGCTGTATGTTGATACACAGTTTCAACCTCACACGAGTGAAGCTTCAGAAATTTG 46714
 178 AAAATTTTAAACTAAATGTTATCATCTCCAAAGACACTATCTCTTAAGTAAATTT 237
 46715 AAAATTTTAAACCAAAATGATATCATCTCCAAAGACACTATCTCTTAAGTAAATTT 46774
 238 TTCCATATTCCTTGGAAATCTTCAGACTTTCTTACTGTGGGCTTGTCCGAGTTGATATGC 297
 46775 TTCCATATTCCTTGGAAATCTTCAGACTTTCTTACTGTGGGCTTGTCCGAGTTGATATGC 46834
 298 GTATGCTTTGCTTAAAGCTTAGCAATTTAGACTTGAAGTGAAGTGAAGTGAAGTGAAGT 357
 46835 GTATGCTTTGCTTAAAGCTTAGCAATTTAGACTTGAAGTGAAGTGAAGTGAAGTGAAGT 46894
 358 TTCCAGCTTACAAATTTGGAGACCTTCATACACCTTCAAGAACTTAACTGAATGACATCACT 417
 46895 TTCCAGCTTACAAATTTGGAGACCTTCATACACCTTCAAGAACTTAACTGAATGACATCACT 46954
 418 TGGAGTCAATTTAGTGTAGCTTGTGTCATTTTACATCTCCAGAACTCACTTCGGAGTTGG 477
 46955 TGGAGTCAATTTAGTGTAGCTTGTGTCATTTTACATCTCCAGAACTCACTTCGGAGTTGG 47314
 478 ACCTTCAGCAAGCAAAATTAAGCACTTCCCTGTGTCAGTTTTCAGCTTCAGCAAGCACTTA 537
 47015 ATCTTCAGCAAGCAAAATTAAGCACTTCCCTGTGTCAGTTTTCAGCTTCAGCAAGCACTTA 47074
 538 AGAATTTTAAACTTGAACATTAATGATTTCAATTTCTTCCGAGATAGCAACTAA 597
 47075 CTAATTTTAAACTTGAACATTAATGATTTCAATTTCTTCCGAGATAGCAACTAA 47134

QY 598 TAAACCTTCCTTTTGTGACAGCTCGAAATAGCTTCCATTTTTCCTAGTAGTAATTA 657
 Db |||||
 47135 CAAACCTTCCTTTTGTGACAGCTCGAAATAGCTTCCATTTTTCCTAGTAGTAATTA 47194
 QY 653 GAAATTTATCCCTTGAATACCTGATCTTTTGGAAATACCTTTTGAACACCAAGTCC 717
 Db |||||
 47195 AATGTTATCCCTTGAATACCTGATCTTTTGGAAATACCTTTTGAACACCAAGTCC 47254
 QY 718 TTCAGTATTAAGCTGCAAGCACCATTAACTTTTATGGAATCTTCTCAGCAACCATAT 777
 Db |||||
 47255 TTCAGTATTAAGCTGCAAGCACCATTAACTTTTATGGAATCTTCTCAGCAACCATAT 47314
 QY 778 TACATTAATAGG 788
 Db |||||
 47315 TATCCTATAGG 47325

RESULT 9

AC099934

LOCUS

DEFINITION

AC099934

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thoman, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkatakrishnan, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (17-DEC-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 17, 2003 this sequence version replaced gi:38502455.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www.seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L13587
Center clone name: L1_K_9

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and the accession number will be preserved.

* 1 163083: contig of 163083 bp in length
* 163084 163183: gap of 100 bp
* 163184 232458: contig of 69275 bp in length.

FEATURES

Source
1. .232458
/location=Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="12"
/map="12"
/clone="RP23-17K9"
/clone_lib="RPCI-23 Female Mouse BAC"

ORIGIN

Query Match 46.6%; Score 566; DB 2; Length 232458;
Best Local Similarity 83.1%; Pred. No. 1.5e-116;
Matches 658; Conservative 0; Mismatches 130; Indels 4; Gaps 1;
QY 1 GAAGAACTAGCATGTATGTTATCTCCAGTGAATTTAATTTCTACAAAC---TTTAA 56
DB 82575 GGAGAACTAGCAAGCATGAATTTAAATGAATTTAATTTGTTAACTGTTTAA 82634
QY 57 TTATTTAGGCAATTTCCAGCAGTTTAAAGGTTCTTTTTCAGCTATGAGACGTGCTAT 116
DB 82635 TTATTTAGGCAATTTCCAGCAGTTTAAAGGTTCTTTTTCAGCTATGAGACGTGCTAT 82694
QY 117 AGAGGCTGTAATGTTGATACACAGTTTCAACGCTCACACGAGTGAAGCTTCAGAAATTT 176
DB 82695 AGAGGCTGTAATGTTGATACACAGTTTCAACGCTCACACGAGTGAAGCTTCAGAAATTT 82754
QY 177 GAAACCTTTAAACTAAATGTTTATCATCTCAAAAGAGCTATCTCTTAAGTAAGAT 236
DB 82755 GAAATATAAAACCAATGTTTATCATCTCAAAAGAGTATCTCTTAAGTAAGAT 82814
QY 237 TTTCATATCTTGGAACTTTTTCAGCTTTCTTCTGAGCTTTCTGAGGCTTTGCGAGTTGATATG 296
DB 82815 TTTCATATTTTCTGGAACACCTTTCAGGCTTTCTTCTGAGCTTTGCGAGTTGATATG 82874
QY 297 CGTATGCTTTGCTTTAAAGGCTTTAGAAATTTAGCTTTGAGTGCACCAACCATATAAAAG 356

Db 82875 CTTATGCTGCTTTAAATAATCTTACGAGTTAGACCTGAGTCTCACTATATAAAAAA 82934
QY 357 CTTCCAGCTCAATTTGGAGACCTCATACACCTTCAAGAACTTTAAACCTGAATGACATCAC 416
DB 82935 CTTCCAGCTCAATTTGGAGACCTCATACACCTTCAAGAACTTTAAACCTGAATGACATCAC 82994
QY 417 TTGGAGCTATTAGTGTAGCTTGTGTCATCTCTACACTCCAGAGTCACTTGGGAGTTTG 476
DB 82995 CTGGAGACCTTTAGTGTGCTTGTGTCATCTCTACACTCCAGAGTCACTTGGGAGTTTG 83054
QY 477 GACCTCAGCAGAAACAAATCAAGCCACTCCCTGTGCAAGTTTTCAGCTCCAGCAACTT 536
DB 83055 GATCTCAGCAGAAACAAATCAAGCCACTCCCTGTGCAAGTTTTCAGCTCCAGCAACTT 83114
QY 537 AAGAAATTTAAACTGACGATAATGAATTTGAATTTTCCTTCAAGATAGGACAACTA 596
DB 83115 ACTAACTTTAAACCTTAATGATAATGAATTTGAATTTTCCTTCAAGATAGGACAACTA 83174
QY 597 ATAAACCTTCGCTTTTGTGAGCAGCTCGAATAAGCTTCCACTTTTGCCTTAGTAATTT 656
DB 83175 ACAACCTTCGCTTTTGTGAGCAGCTCGAATAAGCTTCCACTTTTGCCTTAGTAATTT 83234
QY 657 AGAAATTTATCCCTTGAATATCTTGGATCTTTTGGAAATATCTTTTGAACCAACCAAGTC 716
DB 83235 AAGATGTTATCCCTTGGAGTACTTGGATCTTTTGGAAATATCTTTTGAACCAACCAAGTC 83294
QY 717 CTTCCAGTATAAAGCTGAGCAGCAATTAATTTTGAATTTTGAATTTTGAACCAACCAAGTC 776
DB 83295 ATTCCAATTTAAAGCTTCAAGTACCATAACTTTTGAATTTTGAATTTTGAACCAACCAAGTC 83354
QY 777 TTACATAATAGG 788
DB 83355 CTATCTATAGG 83366

RESULT 10
LOCUS AX255882 682 bp DNA linear PAT 10-OCT-2001
DEFINITION Sequence 33 from Patent WO0170976.
ACCESSION AX255882
VERSION AX255882.1 GI:16074922
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Xu, J., Pyle, R.A. and Stoik, J.A.
TITLE Compositions and methods for the therapy and diagnosis of ovarian and endometrial cancer
JOURNAL Patent: WO 0170976-A 33 27-SEP-2001;
CORIXA CORPORATION (US)
FEATURES
source location/Qualifiers
1. .682
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 42.9%; Score 521.6; DB 6; Length 682;
Best Local Similarity 97.5%; Pred. No. 2.8e-106;
Matches 55; Conservative 0; Mismatches 10; Indels 4; Gaps 2;
QY 277 GGCTTCTCGAGTTGATATGCTGATGCTTGGCTTTAAAGCCTTAGCAATTTAGCTTGA 336
DB 1 GGCTTCTCGAGTTGATATGCTGATGCTTGGCTTTAAAGCCTTAGCAATTTAGCTTGA 60
QY 337 GTCACCAACCATATAAAAGCTTCCAGTACAAATTTGAGACCTTCATACCTTTCAAGAAC 396
DB 61 GTCACCAACCATATAAAAGCTTCCAGTACAAATTTGAGACCTTCATACCTTTCAAGAAC 120
QY 397 TTAACCTGAATGACATGCTGAGTCAATTTAGTGTAGCTTTGTGTCTTACACTCC 456

```

Db      121  TTAACCTGAATGACAACTCAGTTCAGTCAATTTAGTGTAGCTTGTGTGTCATCTTACACTCC 180
QY      457  AGAAGTCACCTTCGAGTTTGGACCTCAGCAGAGACAAATCAAGGCACTCCCTGTGAGT 516
Db      181  AGAAGTCACCTTCGAGTTTGGACCTCAGCAGAGACAAATCAAGGCACTCCCTGTGAGT 240
QY      517  TTTGCCAGCTCCAGGAACCTTAAGAAATTTAAACCTTTGACGATAATGAATTTGATTCAAATTC 576
Db      241  TTTGCCAGCTCCAGGAACCTTAAGAAATTTAAACCTTTGACGATAATGAATTTGATTCAAATTC 302
QY      577  CTTGCAAGATAGGCAACTAATAAACCCTTTCCTTTTGTGACGAGCTGGAATAAGCTTC 636
Db      301  CTTGCAAGATAGGCAACTAATAAACCCTTTCCTTTTGTGACGAGCTGGAATAAGCTTC 360
QY      637  CATTTTGGCTAGTGAATTTAGAAATTTAACCCTTTCCTTTTGTGACGAGCTTTCGAAATA 696
Db      361  CATTTTGGCTAGTGAATTTAGAAATTTAACCCTTTCCTTTTGTGACGAGCTTTCGAAATA 420
QY      697  CTTTGAACAACCAAAAGTCTCCAGTAATAAAGCTGCAAGCACCATTAACTTTATTGG 756
Db      421  CTTTGAACAACCAAAAGTCTCCAGTAATAAAGCTGCAAGCACCATTAACTTTATTGG 480
QY      757  AATCTTCTGCAAGAACCAATATTCATTAAT-AGATTCATATGCTCT---CATATCAT 812
Db      481  AATCTTCTGCAAGAACCAATATTCATTAATTAAGATTCATATGCTCTCTTCATATTCATTT 540
QY      813  CCATTCATCTCTGCCAAGATTGG 837
Db      541  CCATTCATCTCTGCCAAGATTGG 565

RESULT 11
AX336594/c
LOCUS      AX336594      467 bp      DNA      linear      PAT 09-JAN-2002
DEFINITION Sequence 7103 from Patent WO0194629.
ACCESSION AX336594
VERSION   AX336594.1 GI:18127313
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
REFERENCE
AUTHORS   Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE     Cancer gene determination and therapeutic screening using signature
          gene sets
JOURNAL   Patent: WO 0194629-A 7103 13-DEC-2001;
          Avalon Pharmaceuticals (US)
FEATURES
source    1..467
           /organism="Homo sapiens"
           /mol_type="unassigned DNA"
           /db_xref="taxon:9606"

ORIGIN
Query Match 36.5%; Score 443.8; DB 6; Length 467;
Best Local Similarity 98.1%; Pred. No. 7.6e-89;
Matches 460; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

QY      739  CACCAATTAATTTATGGAACTCTCTGCGAGCAACCATATTACATATAGGATTCATATG 798
Db      467  CACCAATTAATTTAT--GGAATCTTGGCAGCAACCATATTACATATAGGATTCATATG 410
QY      799  GCTCTCATATCATTTCCATTCCTCTGCAAGATTTGGATCCGCAAAATTTGTTT 858
Db      409  GCTCTCATATCATTTCCATTCCTCTGCAAGATTTGGATCCGCAAAATTTGTTT 350
QY      859  GTGGAAGATTTGCTGCACTTTTCATTTCAAGGAACCTACTACCATGATTCGATTCG 918
Db      349  GTGGAAGATTTGCTGCACTTTTCATTTCAAGGAACCTACTACCATGATTCGATTCG 290
QY      919  TTGCCACACTGTGCTCTTAGTAGATAATTTGGGTGGTACTGAAGCACCCTATATCTCTT 978

```

```

Db      289  TTGCCACACTGTGCTCTTAGTAGATAATTTGGGTGGTACTGAAGCACCCTATATCTCTT 230
QY      979  AATTCTGTCTCTAGGCTGTTATGTTAATTCCTCTGATATGTTAAAGTAATGGTGAGAC 1038
Db      229  AATTCTGTCTCTAGGCTGTTATGTTAATTCCTCTGATATGTTAAAGTAATGGTGAGAC 170
QY      1039  CAGAAAAGAAATTTCAATTAACAGATCAGTTTGGGTGCTATGATGATTTTGCACGCTCA 1098
Db      169  CAGAAAAGAAATTTCAATTAACAGATCAGTTTGGGTGCTATGATGATTTTGCACGCTCA 110
QY      1099  AATTGGAGTAAAGGAAGATTTCTGTATCTACTTCTGAGAGGAGGAGGAGGAGGAGGAGG 1158
Db      109  AATTGGAGTAAAGGAAGATTTCTGTATCTACTTCTGAGAGGAGGAGGAGGAGGAGGAGG 50
QY      1159  CATTTAGATGATCTCAAAACCTTTTATTTAAACCAATTTTACTTTTAAAA 1207
Db      49  CATTTAGATGATCTCAAAACCTTTTATTTAAACCAATTTTACTTTTAAAA 1

RESULT 12
BC030142
LOCUS      BC030142      990 bp      mRNA      linear      PRI 07-OCT-2003
DEFINITION Homo sapiens peptidylprolyl isomerase (cyclophilin) like 5, mRNA
          (CDNA clone MGC:20689 IMAGE:4764756), complete cds.
ACCESSION BC030142
VERSION   BC030142.2 GI:33988785
KEYWORDS  MGC.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE
AUTHORS   Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Xarusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Udwin, C.B., Toshiyuki, S.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., Loquellano, N.A., Peters, G.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Li, X., Gibbs, R.A.,
Faney, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Smallos, D.E.,
Schnerch, A., Schein, J.E., Jones, S.O. and Marra, M.A.
TITLE     Generation and initial analysis of more than 15,000 full-length
          human and mouse cDNA sequences
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE   22388257
PUBMED    12477932
REFERENCE
AUTHORS   Strausberg, R.
TITLE     Direct Submission
JOURNAL   Submitted (07-MAY-2002) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
REMARK
COMMENT   NIH-MGC Project URL: http://mgc.nci.nih.gov
          On Aug 20, 2003 this sequence version replaced gi:20988783.
          Contact: MGC help desk
          Email: cgabbs@mail.nih.gov
          Tissue Procurement: Louis Staudt
          cDNA Library Preparation: Rubin Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL);
          DNA Sequencing by: National Institutes of Health Intramural
          Sequencing Center (NISC),
          Gaithersburg, Maryland;

```


SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
TITLE Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
JOURNAL Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 6164 09-JUL-2002;
HELEX RESEARCH INSTITUTE
COMMENT PS Homo sapiens (human)
PN JP 2002191363-A/6164
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORI
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10,
C12N5/02, C12N1/08, G06F17/30, C12N15/00, C12N5/00, CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT source 1. 584
FT /organism="Homo sapiens (human)"
FEATURES
source
Location/Qualifiers
1. 584
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 34.5%; Score 418.8; DB 6; Length 584;
Best Local Similarity 99.5%; Pred. No. 2.9e-83;
Matches 420; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 782 TAATAGGATTCCTATGCTCTCATATCCATTCCTGCGCAAGATTTGGATAC 841
Db 422 TATGAGGATTCCTATGCTCTCATATCCATTCCTGCGCAAGATTTGGATAC 363
QY 842 CGCAAAATTTGTTTGTGGAAGATTTCTGCTGAACTCTTTTCAATCAAGGAATCTAC 901
Db 362 CGCAAAATTTGTTTGTGGAAGATTTCTGCTGAACTCTTTTCAATCAAGGAATCTAC 303
QY 902 CATGAATCTGCATTTCTGCGCACTGCTGCTAGTATGATTAATTTGGTGGTACTGA 961
Db 302 CATGAATCTGCATTTCTGCGCACTGCTGCTAGTATGATTAATTTGGTGGTACTGA 243
QY 962 AGCACCATTATCTCTTATTTCTGCTAGGCTGTTATGTTAACTCTCTGATATGTT 1021
Db 242 AGCACCATTATCTCTTATTTCTGCTAGGCTGTTATGTTAACTCTCTGATATGTT 183
QY 1022 AAGTAATGGTGGACACAGAAAGAAATTTCAATACAGATTCAGTTGGGGTGCATGT 1091
Db 182 AAGTAATGGTGGACACAGAAAGAAATTTCAATACAGATTCAGTTGGGGTGCATGT 123
QY 1091 ATGATTTTCAGCGCTCAAAATTTGGAGTAAAGGAATTTCTGTATCTTGTGAGAGAGAG 1141
Db 122 ATGATTTTCAGCGCTCAAAATTTGGAGTAAAGGAATTTCTGTATCTTGTGAGAGAGAG 63
QY 1142 GAATGTGTATGATTCATTTAGATGACCTCCAAACTTTTATTAACCAATTTTAGTT 1201
Db 62 GAATGTGTATGATTCATTTAGATGACCTCCAAACTTTTATTAACCAATTTTAGTT 3
QY 1201 TT 1203
Db 2 TT 1

Sequence 13527 from Patent EP1074617.
ACCESSION AX878622 GI:40033358
VERSION AX878622.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
TITLE Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
JOURNAL Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
Primer for synthesizing full-length cDNA and their use
Patent: EP 1074617-A 13527 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
source
Location/Qualifiers
1. 1590
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
385..882
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE90199.1"
/db_xref="GI:40033358"
/translation="MSVTLHTDVGDIKIEFCERTPKCEMSRCVPOAGVQWRDLGS
LOPPPGFKQVFCLSLPRGTGSGNSINGKKFEDEYSEYLKXRVGVVMMNGPNTNG
SOFETTYGKQPHLDMKTYVFKVIGDLGILDELEKLPVNEKTIPLNDVHLNDITIIHA
NPFAQ"
ORIGIN
Query Match 34.5%; Score 418.8; DB 6; Length 1590;
Best Local Similarity 99.5%; Pred. No. 2.5e-83;
Matches 420; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 782 TAATAGGATTCCTATGCTCTCATATCCATTCCTGCGCAAGATTTGGATAC 841
Db 1169 TATGAGGATTCCTATGCTCTCATATCCATTCCTGCGCAAGATTTGGATAC 1228
QY 842 CGCAAAATTTGTTTGTGGAAGATTTCTGCTGAACTCTTTTCAATCAAGGAATCTAC 901
Db 1229 CGCAAAATTTGTTTGTGGAAGATTTCTGCTGAACTCTTTTCAATCAAGGAATCTAC 1288
QY 902 CATGAATCTGCATTTCTGCGCACTGCTGCTAGTATGATTAATTTGGTGGTACTGA 961
Db 1289 CATGAATCTGCATTTCTGCGCACTGCTGCTAGTATGATTAATTTGGTGGTACTGA 1348
QY 962 AGCACCATTATCTCTTATTTCTGCTAGGCTGTTATGTTAACTCTCTGATATGTT 1021
Db 1349 AGCACCATTATCTCTTATTTCTGCTAGGCTGTTATGTTAACTCTCTGATATGTT 1408
QY 1022 AAGTAATGGTGGACACAGAAAGAAATTTCAATACAGATTCAGTTGGGGTGCATGT 1081
Db 1409 AAGTAATGGTGGACACAGAAAGAAATTTCAATACAGATTCAGTTGGGGTGCATGT 1468
QY 1082 ATGATTTTCAGCGCTCAAAATTTGGAGTAAAGGAATTTCTGTATCTTGTGAGAGAGAG 1141
Db 1469 ATGATTTTCAGCGCTCAAAATTTGGAGTAAAGGAATTTCTGTATCTTGTGAGAGAGAG 1528
QY 1142 GAATGTGTATGATTCATTTAGATGACCTCCAAACTTTTATTAACCAATTTTAGTT 1201
Db 1529 GAATGTGTATGATTCATTTAGATGACCTCCAAACTTTTATTAACCAATTTTAGTT 1588
QY 1201 TT 1203
Db 1589 TT 1590

Search completed: June 16, 2004, 00:15:48
Job time : 3279 secs

RESULT 15
AX878622
LOCUS AX878622 1590 bp DNA linear PAT 17-DEC-2003